

PR	24-NOV-1998;	98US-0109869.	
XX	(INCY-)	INCYTE PHARM INC.	
PA			
XX			
PI	Tang YT, Lal P, Hillman JL, Yue H, Guegler KJ, Corley NC;		
PI	Bandman O, Patterson C, Gorgone GA, Kaser MR, Baughn MR;		
PI	Au-Young J;		
XX			
XX	WPI; 2000-072605/06.		
DR	N-PSDB; AA256761.		
XX			
PT	Proteins, polynucleotides, vectors, host cells and antibodies used to		
PT	diagnose, treat or prevent immune, reproductive, smooth muscle, and		
PT	neurological, gastrointestinal, developmental and cell proliferative		
PT	disorders -		
XX			
PS	Claim 1; Page 163; 229pp; English.		
XX			
CC	AAZ56698 to AAZ56776 encode AAY57877 to AAY57955 which represent human		
CC	transmembrane proteins designated HTMPN-1 to HTMPN-79, respectively.		
CC	The transmembrane protein have immunospecific, antiproliferative and		
XX	neuroprotective activities. The human transmembrane proteins,		
XX	polynucleotides encoding them and other compositions and methods from		
XX	the present invention, can be used for the diagnosis, treatment or		
CC	prevention of immune, reproductive, smooth muscle, neurological,		
CC	gastrointestinal, developmental and cell proliferative disorders. The		
CC	HTMPN's can be used to treat or prevent disorders associated with a		
CC	decreased expression or activity of HTMPN.		
XX			
SQ	Sequence 129 AA;		
	Query Match 93.9%; Score 571; DB 21; Length 129;		
	Best Local Similarity 94.7%; Pred. No. 1.8e-52;		
	Matches 107; Conservative 0; Mismatches 6; Indels 0; Gaps		
Oy	1 MARGSLRLLRLVLVLGLWALLRSVAGEQAGTAPCSRGSWSADLDKCMDCASCRRPH 60		
Db			
	1 margslrrllrlvlvlgllwllrrsvageqagtapcsggsswsadldkcmdcascrrph 60		
Oy	61 SDFCLGCAAPAPAPFLLWPILGGALSLTFVLGLLSGFLWRRCRRSPP 113		
Db			
	61 sdfclgcaapapfllwpilggalsltfvlglsgflwrrcrrrefttp 113		
RESULT	4		
AAU03498			
ID	AAU03498 standard; Protein; 129 AA.		
XX			
	AAU03498;		
XX			
XX	26-SEP-2001 (first entry)		
DE	Human TWEAK receptor (TWEAKR) polypeptide.		
XX			
KW	TWEAK receptor; TWEAKR; tumour necrosis factor; TNF; angiogenesis;		
KW	ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;		
KW	retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;		
KW	rubeosis; uveitis; macular degeneration; arthritis; rheumatism;		
KW	corneal graft neovascularisation; psoriasis; metastatic condition;		
KW	malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;		
KW	preneoplastic condition; myocardial angiogenesis; wound granulation;		
KW	scleroderma; vascular adhesion; telangiectasia; ischaemia; human;		
KW	atherosclerotic plaque neovascularisation; coronary atherosclerosis;		
KW	peripheral atherosclerosis.		
XX			
OS	Homo sapiens.		
XX			
XX			
FT	Key Location/Qualifiers		
FT	Peptide 1..27		
FT	/note= "Signal peptide"		
FT	28..129		
FT	/note= "Mature human TWEAKR protein"		
FT			

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2002, 09:55:03; Search time 23.81 Seconds
(without alignments)
354.656 Million cell updates/sec

Title: US-09-690-454-59
Perfect score: 608
Sequence: 1 MARGSLRLLVLLGLWLA.....LSGFLVRRRRSSPPPX 114

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

A_Geneseq_1101.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	607	99.8	114	AAW73409	Human secreted pro
2	571	93.9	129	AAW88506	Human liver clone
3	571	93.9	129	AAV57940	Human transmembran
4	571	93.9	129	AAU03498	Human TWEAK recept
5	433	71.2	309	AAU03500	Human TWEAK recept
6	274.5	45.1	112	AAV91463	Human secreted pro
7	274.5	45.1	155	AAV91604	Human secreted pro
8	274.5	45.1	156	AAV91552	Human secreted pro
9	80.5	13.2	426	AAV01337	TNF receptor apopt
10	80.5	13.2	508	AAV80631	2B10 AMH-receptor.
11	79	13.0	224	AAV95174	Human protein sequ

12	77	12.7	448	21	AAV56750	Smooth muscle prol
13	77	12.7	448	21	AAV54990	Full length mouse
14	76.5	12.6	250	21	AAV58208	Canine mature Fit-
15	76.5	12.6	276	21	AAV58207	Canine Fit-3 ligan
16	74	12.2	234	22	AAW24485	Human EST encoded
17	73.5	12.1	187	20	AAV29189	Amino acid sequenc
18	73	12.0	141	20	AAV73883	Human prostate tum
19	73	12.0	141	20	AAV59770	Human normal ovari
20	72.5	11.9	242	22	AAV92250	C glutamicum prote
21	72.5	11.9	365	20	AAV43138	Human NAIL protein
22	72.5	11.9	365	22	AAW23942	Human EST encoded
23	72	11.8	1278	20	AAW88445	Human NPC1 (Nieman
24	72	11.8	1278	21	AAW42983	Human ORFX ORF2747
25	72	11.8	3724	18	AAW23718	Platenolide syntha
26	72	11.8	3724	18	AAW22608	Platenolide syntha
27	71.5	11.8	381	19	AAW81597	Protein encoded by
28	71.5	11.8	2963	19	AAW56444	Fragment HGF1789 o
29	71	11.7	308	22	AAW41162	Human polypeptide
30	71	11.7	349	22	AAW42035	Human polypeptide
31	70	11.5	301	20	AAV13941	Human transmembran
32	70	11.5	301	22	AAW73498	Human transferase
33	70	11.5	305	22	AAW39376	Human polypeptide
34	70	11.5	443	21	AAW84706	Amino acid sequenc
35	70	11.5	461	21	AAV56752	Smooth muscle prol
36	70	11.5	461	21	AAV54991	Full length mouse
37	70	11.5	531	21	AAV84361	Amino acid sequenc
38	70	11.5	567	22	AAW70531	Human Prol protein
39	70	11.5	720	21	AAW88280	Membrane-bound pro
40	70	11.5	720	21	AAV66695	Human TANGO 215 pr
41	70	11.5	720	22	AAU00401	Human secreted pro
42	70	11.5	720	22	AAW87544	Human Prol1344, Ho
43	70	11.5	720	22	AAW70532	Human Prol2 protein
44	70	11.5	720	22	AAW65218	Human Prol1344 (UNQ
45	70	11.5	737	22	AAW93670	Human protein sequ

ALIGNMENTS

RESULT 1

AAW73409
ID AAW73409 standard; Protein; 114 AA.

XX AAW73409;

XX 19-FEB-1999 (first entry)

DT Human secreted protein encoded by Gene No. 13.

DE Secreted protein; human; protein therapy; gene therapy; blood disorder;
pathological condition; diagnosis; cancer; neurological disorder;
developmental abnormality; foetal deficiency; leukaemia; hepatic disease;
immune system disorder; Alzheimer's disease; cognitive disorder;
schizophrenia; prostate disease; autoimmune disorder; AIDS.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 114 /note= "unspecified amino acid"

XX WO9854206-A1.

XX 03-DEC-1998.

XX 28-MAY-1998; 98WO-US10868.

XX 29-AUG-1997; 97US-0056296.

XX 30-MAY-1997; 97US-0044039.

XX 30-MAY-1997; 97US-0048093.

XX 30-MAY-1997; 97US-0048101.

XX 30-MAY-1997; 97US-0048190.

XX 30-MAY-1997; 97US-0048356.


```

PR 24-NOV-1998; 98US-0109869.
XX (INCY-) INCYTE PHARM INC.
XX
XX Tang YT, Lal P, Hillman JL, Yue H, Guegler KJ, Corley NC;
XX Bandman O, Patterson C, Gorgone GA, Kaser MR, Baughn MR;
XX Au-Young J;
XX
XX WPI; 2000-072605/06.
XX N-PSDB; AA256761.
XX
XX Proteins, polynucleotides, vectors, host cells and antibodies used to
XX diagnose, treat or prevent immune, reproductive, smooth muscle,
XX neurological, gastrointestinal, developmental and cell proliferative
XX disorders -
XX
XX Claim 1; Page 163; 229pp; English.
XX
XX AAZ56698 to AAZ56776 encode AAY57877 to AAY57955 which represent human
XX transmembrane proteins designated HTPNP-1 to HTPNP-79, respectively.
XX The transmembrane protein have immunospecific, antiproliferative and
XX neuroprotective activities. The human transmembrane proteins,
XX polynucleotides encoding them and other compositions and methods from
XX the present invention, can be used for the diagnosis, treatment or
XX prevention of immune, reproductive, smooth muscle, neurological,
XX gastrointestinal, developmental and cell proliferative disorders. The
XX HTPNP's can be used to treat or prevent disorders associated with a
XX decreased expression or activity of HTPNP.
XX
XX Sequence 129 AA;
SQ
- Query Match 93.9%; Score 571; DB 21; Length 129;
Best Local Similarity 94.7%; Pred. No. 1.8e-52;
Matches 107; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 MARGSLRLLRLVGLWALLRSVAGQAPGTAPCRGSSWSADLDKCMDCASCARHPH 60
Db 1 margslrrllrlvlgllwllrsvagqapgtapcrgsswsadldkcmdcascarp 60
QY 61 SDFCLGCAAAPPPAPFRLWPILGGALSITFVLGSLGFLVWRRCRRSSPPP 113
Db 61 sdfclgcaaaappapfrrllwpilggalsitfvllgslgflvrrrrrckfttp 113
RESULT 4
AAU03498
ID AAU03498 standard; Protein; 129 AA.
AC
XX AAU03498;
XX
XX 26-SEP-2001 (first entry)
XX
XX Human TWEAK receptor (TWEAKR) polypeptide.
XX
XX TWEAK receptor; TWEAKR; tumour necrosis factor; TNF; angiogenesis;
XX ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;
XX retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;
XX rubeosis; uveitis; macular degeneration; arthritis; rheumatism;
XX corneal graft neovascularisation; psoriasis; metastatic condition;
XX malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;
XX preneoplastic condition; myocardial angiogenesis; wound granulation;
XX scleroderma; vascular adhesion; telangiectasia; ischaemia; human;
XX atherosclerotic plaque neovascularisation; coronary atherosclerosis;
XX peripheral atherosclerosis.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..27
XX FT /note= "Signal peptide"
XX Protein 28..129
XX FT /note= "Mature human TWEAKR protein"

```

```

FT Domain 1..78 "Extracellular domain"
FT FT 79..101
FT Domain "Transmembrane domain"
FT FT 102..129
FT FT /note= "Intracellular domain"
XX
XX WO200145730-A2.
XX
XX 28-JUN-2001.
XX
XX 19-DEC-2000; 2000WO-US34755.
XX
XX 20-DEC-1999; 99US-0172878.
XX 10-MAY-2000; 2000US-0203347.
XX
XX (IMMV ) IMMUNEX CORP.
XX
XX WILEY SR;
XX
XX WPI; 2001-417975/44.
XX N-PSDB; AAS03563.
XX
XX Modulating angiogenesis in a mammal for treating diseases mediated by
XX angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or
XX peripheral tissue, by administering antagonist or agonist of TWEAK
XX receptor -
XX
XX Example 1; Fig 1; 46pp; English.
XX
XX The sequence represents the human TWEAK receptor (TWEAKR) protein. The
XX TWEAK protein is a member of the tumour necrosis factor (TNF) family and
XX induces angiogenesis. TWEAKR may therefore be used to screen for and
XX develop TWEAKR agonists and antagonists for the modulation of
XX angiogenesis, to be used in the treatment and diagnosis of human disease.
XX The disorders mediated by angiogenesis include ocular disorders
XX characterised by ocular neovascularisation such as diabetic retinopathy,
XX neovascular glaucoma, retinoblastoma, retinopathy of prematurity,
XX retrolental fibroplasia, rubeosis, uveitis, macular degeneration and
XX corneal graft neovascularisation, and inflammatory diseases such as
XX arthritis, rheumatism and psoriasis. Other treatable diseases include
XX malignant and metastatic conditions such as sarcomas and carcinomas,
XX benign tumours and preneoplastic conditions, myocardial angiogenesis,
XX haemophilic joints, scleroderma, vascular adhesions, atherosclerotic
XX plaque neovascularisation, telangiectasia, wound granulation, coronary
XX atherosclerosis, peripheral atherosclerosis and ischaemia.
XX
XX Sequence 129 AA;
SQ
Query Match 93.9%; Score 571; DB 22; Length 129;
Best Local Similarity 94.7%; Pred. No. 1.8e-52;
Matches 107; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 MARGSLRLLRLVGLWALLRSVAGQAPGTAPCRGSSWSADLDKCMDCASCARHPH 60
Db 1 margslrrllrlvlgllwllrsvagqapgtapcrgsswsadldkcmdcascarp 60
QY 61 SDFCLGCAAAPPPAPFRLWPILGGALSITFVLGSLGFLVWRRCRRSSPPP 113
Db 61 sdfclgcaaaappapfrrllwpilggalsitfvllgslgflvrrrrrckfttp 113
RESULT 5
AAU03500
ID AAU03500 standard; Protein; 309 AA.
XX
XX AAU03500;
XX
XX 26-SEP-2001 (first entry)
XX
XX Human TWEAK receptor-Fc (TWEAKR-Fc) fusion polypeptide.
XX

```


CC cancer, tumours, neurodegenerative disorders, developmental abnormalities
 CC and foetal deficiencies, blood disorders, diseases of the immune system,
 CC autoimmune diseases, hepatic and renal disease, inflammation,
 CC allergies, Alzheimer's and behavioural disorders, schizophrenia,
 CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
 CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
 CC cardiovascular disorders, reproductive disorders, gastrointestinal
 CC disorders, respiratory disorders and metabolic disorders. The proteins
 CC or polynucleotides can also be used as food additives or preservatives.
 CC The proteins are also useful for identifying their binding partners.
 CC AAA26337 to AAA26345 and AAY91450 are sequences used in the
 CC exemplification of the present invention.
 XX
 SQ Sequence 112 AA;

Query Match 45.1%; Score 274.5; DB 21; Length 112;
 Best Local Similarity 96.4%; Pred. No. 1.9e-21;
 Matches 54; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
 QY 1 MARGSLRLRLVLGLWALLRSVAGEQAPGTAPCSRGSSWSADLCKMDCSC-SC 55
 DB 1 margslrrllrlvlglwallrsvageqapgtapcsgsswsadldkcmdcstsc 56

RESULT 7
 AAY91604
 ID AAY91604 standard; Protein; 155 AA.
 XX
 AC AAY91604;
 XX
 DT 29-JUN-2000 (first entry)
 XX
 DE Human secreted protein sequence encoded by gene 13 SEQ ID NO:277.
 KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
 KW antiHIV; antiinflammatory; nontropic; neuroprotective; antiallergic;
 KW osteopathic; antiarthritic; antibacterial; antidiabetic; antiasthma;
 KW antipsoriatic; cardiant; gene therapy; cancer; neurological disorder;
 KW immune disease; inflammation; blood disorder; tumour; chromosome 16.
 XX
 OS Homo sapiens.
 XX
 PN WO200006698-A1.
 XX
 PD 10-FEB-2000.
 XX
 PF 29-JUL-1999; 99WO-US17130.
 XX
 PR 30-JUL-1998; 98US-0094657.
 PR 05-AUG-1998; 98US-0095486.
 PR 06-AUG-1998; 98US-0095454.
 PR 06-AUG-1998; 98US-0095455.
 PR 12-AUG-1998; 98US-0096319.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Komatsoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;
 PI Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;
 PI Soppet DR, Endress GA, Ebner R, Olsen HS, Mucenski M;
 XX
 DR WPI; 2000-195282/17.
 XX
 PT New isolated human genes and the secreted polypeptides they encode,
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders -
 XX
 PS Disclosure; Page 36-37; 634pp; English.
 XX
 CC The polynucleotide sequences given in AAA26346 to AAA26458 encode the
 CC human secreted proteins given in AAY91451 to AAY91691. The human secreted
 CC proteins can have activities based on the tissues and cells they are
 CC expressed in. Examples of the activities are: cytostatic;

CC immunosuppressive; antiHIV; antiinflammatory; nontropic; neuroprotective;
 CC antiallergic; osteopathic; antiarthritic; antibacterial; antidiabetic;
 CC antiasthma; antipsoriatic; and cardiant. The polynucleotides and their
 CC corresponding secreted proteins are useful for preventing, treating or
 CC ameliorating medical conditions, e.g. by protein or gene therapy. Also
 CC pathological conditions can be diagnosed by determining the amount of the
 CC proteins in a sample or by determining the presence of mutations in the
 CC polynucleotides. Specific uses are described for each of the
 CC polynucleotides, based on which tissues they are most highly expressed
 CC in, and include developing products for the diagnosis or treatment of
 CC cancer, tumours, neurodegenerative disorders, developmental abnormalities
 CC and foetal deficiencies, hepatic and renal disease, inflammation,
 CC autoimmune diseases, AIDS, sepsis, acne, psoriasis,
 CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
 CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
 CC cardiovascular disorders, reproductive disorders, gastrointestinal
 CC disorders, respiratory disorders and metabolic disorders. The proteins
 CC or polynucleotides can also be used as food additives or preservatives.
 CC The proteins are also useful for identifying their binding partners.
 CC AAA26337 to AAA26345 and AAY91450 are sequences used in the
 CC exemplification of the present invention.
 XX
 SQ Sequence 155 AA;

Query Match 45.1%; Score 274.5; DB 21; Length 155;
 Best Local Similarity 96.4%; Pred. No. 2.7e-21;
 Matches 54; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
 QY 1 MARGSLRLRLVLGLWALLRSVAGEQAPGTAPCSRGSSWSADLCKMDCSC-SC 55
 DB 1 margslrrllrlvlglwallrsvageqapgtapcsgsswsadldkcmdcstsc 56

RESULT 8
 AAY91552
 ID AAY91552 standard; Protein; 156 AA.
 XX
 AC AAY91552;
 XX
 DT 29-JUN-2000 (first entry)
 XX
 DE Human secreted protein sequence encoded by gene 13 SEQ ID NO:225.
 XX
 KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
 KW antiHIV; antiinflammatory; nontropic; neuroprotective; antiallergic;
 KW osteopathic; antiarthritic; antibacterial; antidiabetic; antiasthma;
 KW antipsoriatic; cardiant; gene therapy; cancer; neurological disorder;
 KW immune disease; inflammation; blood disorder; tumour.
 XX
 OS Homo sapiens.
 XX
 PN WO200006698-A1.
 XX
 PD 10-FEB-2000.
 XX
 PF 29-JUL-1999; 99WO-US17130.
 XX
 PR 30-JUL-1998; 98US-0094657.
 PR 05-AUG-1998; 98US-0095486.
 PR 06-AUG-1998; 98US-0095454.
 PR 06-AUG-1998; 98US-0095455.
 PR 12-AUG-1998; 98US-0096319.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Komatsoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;
 PI Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;
 PI Soppet DR, Endress GA, Ebner R, Olsen HS, Mucenski M;
 XX
 DR WPI; 2000-195282/17.
 DR N-PSDB; AAA26447.

XX New isolated human genes and the secreted polypeptides they encode,
PT useful for diagnosis and treatment of e.g. cancers, neurological
XX disorders, immune diseases, inflammation or blood disorders -
XX
PS Claim 11; Page 528; 634pp; English.
XX
CC The polynucleotide sequences given in AAA26346 to AAA26458 encode the
CC human secreted proteins given in AAY91451 to AAY91691. The human secreted
CC proteins can have activities based on the tissues and cells they are
CC expressed in. Examples of the activities are: cytostatic;
CC immunosuppressive; antiHIV; antiinflammatory; neurotropic; neuroprotective;
CC antiallergic; osteopathic; antiarthritic; antibacterial; antidiabetic;
CC antiasthma; antipsoriatic; and cardiant. The polynucleotides and their
CC corresponding secreted proteins are useful for preventing, treating or
CC ameliorating medical conditions, e.g. by protein or gene therapy. Also
CC pathologic conditions can be diagnosed by determining the amount of the
CC proteins in a sample or by determining the presence of mutations in the
CC polynucleotides. Specific uses are described for each of the
CC polynucleotides, based on which tissues they are most highly expressed
CC in, and include developing products for the diagnosis or treatment of
CC cancer, tumours, neurodegenerative disorders, developmental abnormalities
CC and foetal deficiencies, blood disorders, diseases of the immune system,
CC autoimmune diseases, hepatic and renal disease, inflammation,
CC allergies, Alzheimer's and behavioural disorders, schizophrenia,
CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
CC cardiovascular disorders, reproductive disorders, gastrointestinal
CC disorders, respiratory disorders and metabolic disorders. The proteins
CC or polynucleotides can also be used as food additives or preservatives.
CC The proteins are also useful for identifying their binding partners.
CC AAA26337 to AAA26345 and AAY91450 are sequences used in the
CC exemplification of the present invention.
XX
SQ Sequence 156 AA;

Query Match 45.1%; Score 274.5; DB 21; Length 156;
Best Local Similarity 96.4%; Pred. No. 2.7e-21;
Matches 54; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

OY 1 MARGSLRLRLVGLWLLRSVAGEQAPGTAPCSRGSSWSADLDKCMDCSCA-SC 55
Db 1 margslrlrlvlgllwllrsvagedapgtapcsgsswsadldkcmcdstsc 56

RESULT 9
AAB01337
ID AAB01337 standard; Protein; 426 AA.
XX AC AAB01337;
XX
XX DT 25-SEP-2000 (first entry)
XX
XX DE TNF receptor apoptosis mediated protein (TRAMP) death receptor.
XX
XX UL144; death receptor; apoptosis; programmed cell death; FAS;
KW TNF-R1; TRAMP; DR-6; TRAIL; modulation; treatment; cancer; virus;
KW human.
XX
XX OS Homo sapiens.
XX
XX PN WO200034335-A2.
XX
XX PD 15-JUN-2000.
XX
XX PF 03-DEC-1999; 99WO-US26035.
XX
XX PR 04-DEC-1998; 98US-0205018.
XX
XX PA (SCHE) SCHERING CORP.
XX
XX PI Leong C, Phillips JH;

XX WPI: 2000-423383/36.
XX
XX Purified or recombinant polypeptide for modulating apoptosis comprises
PT a sequence which binds to an antibody specific for ULL144 or its
PT fragments
XX
XX Disclosure: Page 64-65; 76pp; English.
XX
XX A pure or recombinant polypeptide which binds to a polyclonal antibody
CC specific for the mature ULL144 is useful for screening molecules which
CC block induction of apoptosis or interfere with antiapoptotic activity.
CC The polypeptide is also useful for modulating apoptosis and useful in
CC treatment of conditions associated with abnormal physiology or
CC development, such as cancer or degenerative conditions and for
CC regulation of viral infection and replication. At least five
CC different death receptors are known, which include the CD95
CC (Fas/APO-1), the TNF receptor-1, TNF receptor apoptosis-mediated
CC protein (TRAMP), death receptor-6 (DR-6), and TNF-related
CC apoptosis-inducing ligand (TRAIL) receptors 1, 2 and 4.
XX
SQ Sequence 426 AA;

Query Match 13.2%; Score 80.5; DB 21; Length 426;
Best Local Similarity 22.2%; Pred. No. 1.7;
Matches 32; Conservative 13; Mismatches 34; Indels 65; Gaps 6;

OY 4 GSLRRLRLVGLWLLRSVAGEQAPGTAPCSRGSSWSADLDKCMDCSCRA--RPHS 61
Db 145 galhrhrlll-----csrrdt-----dcgtclpgfyehg 173
OY 62 DFCGLCAAAPAPFRLW-----PILGALSLSLTVLGLLSGLV-----W 101
Db 174 dgcvsctpppslagavqsvpagsvgrvfwvqvlaglvvpillgatltty 233
OY 102 RRC-----RRRSPPPP 113
Db 234 rhcwphkplvtadeagmealtppp 257

RESULT 10
AAR80631
ID AAR80631 standard; Protein; 508 AA.
XX AC AAR80631;
XX
XX DT 29-FEB-1996 (first entry)
XX
XX DE 2B10 AMH-receptor.
XX
XX 2B10 anti-Mullerian hormone receptor; anti-Mullerian hormone; AMH;
KW receptor; antibody; therapy; H1; 2B10; tumour;
KW transforming growth factor-beta; TGF-beta; bone morphogenic protein;
KW activin; Mullerian inhibiting substance; MIS.
XX
XX OS Oryctolagus cuniculus.
XX

Key Location/Qualifiers
FH Peptide 1..17
FT /note= "signal peptide"
FT Protein 18..508
FT /note= "mature protein"
FT Domain 18..83
FT /note= "extracellular domain"
FT Misc-difference 58..60
FT /note= "potential N-linked glycosylation site"
FT Region 84..109
FT /note= "transmembrane region"
FT Domain 110..508
FT /note= "cytoplasmic domain"
FT Misc-difference 18..508
FT /note= "claimed"

DT 22-FEB-2000 (first entry)
 XX Smooth muscle proliferation modulating protein.
 DE
 XX
 XX Cell proliferation; vascular; smooth muscle cell; arteriosclerosis; PTCA;
 KW endothelial thickening; percutaneous transluminal coronary angioplasty;
 KW myoma; hematopoietic cell-regulation; cytokine; tissue generation; taxis;
 KW actin activity; chemotaxis; blood coagulation; thrombotic; tumor; mouse;
 KW metastasis; nutrient.
 XX
 XX Mus musculus.
 OS
 XX WO955863-A1.
 PN
 XX 04-NOV-1999.
 PD
 XX 28-APR-1999; 99WO-JP02283.
 PF
 XX 28-APR-1998; 98JP-0119731.
 PR
 XX (ONJOY) ONO PHARM CO LTD.
 PA
 XX Honjo T, Tashiro K, Nakamura T;
 PI WPI; 2000-038646/03.
 DR N-PSDB; AAZ39383, AAZ39384.
 DR
 XX Novel mouse polypeptides for treatment, e.g. arteriosclerosis and myoma
 PT
 XX
 XX Claim 1; Page 42-44; 70pp; Japanese.
 XX
 XX The invention provides mouse polypeptides for treatment of diseases due
 CC to abnormal proliferation of smooth muscle. The polypeptides can be
 CC produced by standard recombinant methodology. The polypeptides can be
 CC used according to their inhibition of the proliferation of vascular
 CC smooth muscle cells, particularly in treating arteriosclerosis or re-
 CC narrowing by vascular endothelial thickening after percutaneous
 CC transluminal coronary angioplasty (PTCA), or myoma, hematopoietic cell-
 CC regulatory activity, cytokine activity, tissue generation/repair
 CC activity, actin/inhibitor activity, taxis and chemotaxis activity, blood
 CC coagulation/thrombotic activity, receptor/ligand activity, cadherin/
 CC tumour metastasis inhibiting activity; tumor inhibition, and as nutrient.
 CC The present sequence represents the protein of the invention which can be
 CC used for modulating smooth muscle cell proliferation.
 XX
 XX Sequence 448 AA;
 SQ
 Query Match 12.7%; Score 77; DB 21; Length 448;
 Best Local Similarity 24.1%; Pred. No. 4.3;
 Matches 26; Conservative 8; Mismatches 24; Indels 50; Gaps 5;
 QY 6 LRRLLRLVLGLLALLRSVAGEAPGTA--PCSRGSSWSADLDKCMDCASCRRARPHSDF 63
 Db 4 lkriltvtlailwl-----phpgnaqqgctngfdldrgsgqclidectrtipea-- 52
 QY 64 CLG-----C-----AAAPPAP 74
 Db 53 crgdmcmvqngqgylcprtntpyrgpysnpystsysgypaaappvp 100
 RESULT 13
 AAY54990
 ID AAY54990 standard; Protein; 448 AA.
 XX
 XX AC AAY54990;
 XX
 DT 15-FEB-2000 (first entry)
 DE Full length mouse A55 protein sequence.
 XX
 XX A55 protein; mouse; smooth muscle proliferation; tissue generation;

KW vascular smooth muscle cell; arteriosclerosis; tissue reparation; myoma;
 KW vascular endothelial thickening; hematopoietic cell-regulator; cytokine;
 KW percutaneous transluminal coronary angioplasty; blood coagulation; PTCA;
 KW actin; inhibitor; chemotaxis; thrombosis; cadherin; therapy;
 KW tumour metastasis inhibitor.
 XX
 XX Mus musculus.
 OS
 XX WO955864-A1.
 PN
 XX 04-NOV-1999.
 PD
 XX 28-APR-1999; 99WO-JP02284.
 PF
 XX 28-APR-1998; 98JP-0119731.
 PR
 XX (ONJOY) ONO PHARM CO LTD.
 PA
 XX Honjo T, Tashiro K, Nakamura T;
 PI WPI; 2000-038647/03.
 DR N-PSDB; AAZ40029.
 DR
 XX Novel human polypeptides for treatment of, e.g. arteriosclerosis and
 PT myoma -
 XX
 XX Example 4; Page 52-55; 87pp; Japanese.
 XX
 XX This sequence is the mouse A55 protein. The invention relates to the
 CC human A55 protein. The protein can be used for the treatment of diseases
 CC due to abnormal proliferation of smooth muscle. The polypeptides can be
 CC used according to their inhibition of the proliferation of vascular smooth
 CC muscle cells, particularly in treating arteriosclerosis or re-narrowing
 CC by vascular endothelial thickening after percutaneous transluminal
 CC coronary angioplasty (PTCA), or myoma, hematopoietic cell-regulatory
 CC activity, cytokine activity, tissue generation/repair activity,
 CC actin/inhibitor activity, taxis and chemotaxis activity, blood
 CC coagulation/thrombotic activity, receptor/ligand activity,
 CC cadherin/tumour metastasis inhibiting activity; tumour inhibition, and as
 CC a nutrient.
 XX
 XX Sequence 448 AA;
 SQ
 Query Match 12.7%; Score 77; DB 21; Length 448;
 Best Local Similarity 24.1%; Pred. No. 4.3;
 Matches 26; Conservative 8; Mismatches 24; Indels 50; Gaps 5;
 QY 6 LRRLLRLVLGLLALLRSVAGEAPGTA--PCSRGSSWSADLDKCMDCASCRRARPHSDF 63
 Db 4 lkriltvtlailwl-----phpgnaqqgctngfdldrgsgqclidectrtipea-- 52
 QY 64 CLG-----C-----AAAPPAP 74
 Db 53 crgdmcmvqngqgylcprtntpyrgpysnpystsysgypaaappvp 100
 RESULT 14
 AAY58208
 ID AAY58208 standard; Protein; 250 AA.
 XX
 XX AC AAY58208;
 XX
 DT 14-MAR-2000 (first entry)
 DE Canine mature Flt-3 ligand splice variant clone 1.
 XX
 KW Flt-3 ligand; antibody; canine; feline; inhibitor; immune response;
 KW immunoregulation; tumour; cancer; autoimmune disease; vaccine.
 XX
 XX Canis familiaris.
 OS
 XX WO9961618-A2.
 PN

XX PD 02-DEC-1999.
 XX PF 28-MAY-1999; 99WO-US11942.
 XX PR 29-MAY-1998; 98US-0087306.
 XX PA (HESK-) HESKA CORP.
 XX PI Sim G, Yang S, Dreitz MJ, Wonderling RS;
 XX WPI: 2000-072623/06.
 XX DR N-PSDB; AAZ55507, AAZ55508.
 XX PT Nucleic acids encoding immunoregulatory proteins from cats or dogs,
 XX PT useful for treating or preventing e.g. tumors or autoimmune disease.
 XX PS Claim 3b; Page 177-178; 264pp; English.
 XX Sequences AAY58204 and AAY58206-Y58209 respectively represent
 CC encoded, mature, clone 1 and clone 19 canine Flt-3 ligand, while
 CC sequences AAY58210-Y58211 represent encoded and mature feline Flt-3
 CC ligand. The invention relates to canine interleukin-4 (IL-4),
 CC canine or feline Flt-3 ligand, canine or feline CD40, canine or
 CC feline CD154 (CD40 ligand), canine IL-5, canine IL-13, feline
 CC interferon-alpha (IFN-alpha) and feline granulocyte macrophage
 CC colony-stimulating factor (GM-CSF), and nucleotides which encode these
 CC immunoregulatory proteins. The proteins, their associated
 CC nucleic acids, specific antibodies and inhibitors may be used as
 CC vaccines for therapeutic or prophylactic regulation of an immune
 CC response in animals (particularly cats, dogs, horses and humans).
 CC They may be used to treat autoimmune or infectious diseases including
 CC allergies, tumours, inflammation and graft rejection, and to increase
 CC the response from a co-administered antigen. The nucleotide sequences
 CC can also be used for the recombinant production of a protein, while
 CC nucleotide fragments are useful as probes, as amplification primers and
 CC as sources of inhibitory therapeutics (e.g., antisense
 CC oligonucleotides). The proteins may be used to raise antibodies and to
 CC screen for modulators of activity, while the antibodies may be used in
 CC detection, and in drug targeting.
 XX SQ Sequence 250 AA;

Query Match 12.6%; Score 76.5; DB 21; Length 250;
 Best Local Similarity 27.5%; Pred. No. 2.6;
 Matches 36; Conservative 15; Mismatches 47; Indels 33; Gaps 5;

QY 11 RLVLGLWLALLRSVAG-----EQAPGTAPCSRGSWSA 44
 Db 49 rlvlaqrwmvrlqavagsgmqilleavnteihvftcfaqgtsqqlaalkpwitrnfsq 108
 QY 45 DLD-KCMDCASCARPHSDFLGCAA--APPAPFRLLPILGALSFTFVLGLSGFLVW 101
 Db 109 clclqcpdsstlvpprspgaletalpapqap-rlllll---lpvallmstawclhw 164
 QY 102 RRCRRERSPP 112
 Db 165 rrrrrrrsyp 175

RESULT 15
 AAY58207
 ID AAY58207 standard; Protein; 276 AA.
 XX AC
 XX AAY58207;

XX 14-MAR-2000 (first entry)

XX Canine Flt-3 ligand splice variant clone 1.

XX DE Flt-3 ligand; antibody; canine; feline; inhibitor; immune response;
 XX KW immunoregulation; tumour; cancer; autoimmune disease; vaccine.

XX OS Canis familiaris.
 XX PN WO9961618-A2.
 XX PD 02-DEC-1999.
 XX PF 28-MAY-1999; 99WO-US11942.
 XX PR 29-MAY-1998; 98US-0087306.
 XX PA (HESK-) HESKA CORP.
 XX PI Sim G, Yang S, Dreitz MJ, Wonderling RS;
 XX WPI: 2000-072623/06.
 XX DR N-PSDB; AAZ55503, AAZ55504, AAZ55505, AAZ55506.
 XX PT Nucleic acids encoding immunoregulatory proteins from cats or dogs,
 XX PT useful for treating or preventing e.g. tumors or autoimmune disease
 XX PS Claim 3b; Page 172-173; 264pp; English.
 XX Sequences AAY58204 and AAY58206-Y58209 respectively represent
 CC encoded, mature, clone 1 and clone 19 canine Flt-3 ligand, while
 CC sequences AAY58210-Y58211 represent encoded and mature feline Flt-3
 CC ligand. The invention relates to canine interleukin-4 (IL-4),
 CC canine or feline Flt-3 ligand, canine or feline CD40, canine or
 CC feline CD154 (CD40 ligand), canine IL-5, canine IL-13, feline
 CC interferon-alpha (IFN-alpha) and feline granulocyte macrophage
 CC colony-stimulating factor (GM-CSF), and nucleotides which encode these
 CC immunoregulatory proteins. The proteins, their associated
 CC nucleic acids, specific antibodies and inhibitors may be used as
 CC vaccines for therapeutic or prophylactic regulation of an immune
 CC response in animals (particularly cats, dogs, horses and humans).
 CC They may be used to treat autoimmune or infectious diseases including
 CC allergies, tumours, inflammation and graft rejection, and to increase
 CC the response from a co-administered antigen. The nucleotide sequences
 CC can also be used for the recombinant production of a protein, while
 CC nucleotide fragments are useful as probes, as amplification primers and
 CC as sources of inhibitory therapeutics (e.g., antisense
 CC oligonucleotides). The proteins may be used to raise antibodies and to
 CC screen for modulators of activity, while the antibodies may be used in
 CC detection, and in drug targeting.
 XX SQ Sequence 276 AA;

Query Match 12.6%; Score 76.5; DB 21; Length 276;
 Best Local Similarity 27.5%; Pred. No. 2.9;
 Matches 36; Conservative 15; Mismatches 47; Indels 33; Gaps 5;

QY 11 RLVLGLWLALLRSVAG-----EQAPGTAPCSRGSWSA 44
 Db 75 rlvlaqrwmvrlqavagsgmqilleavnteihvftcfaqgtsqqlaalkpwitrnfsq 134
 QY 45 DLD-KCMDCASCARPHSDFLGCAA--APPAPFRLLPILGALSFTFVLGLSGFLVW 101
 Db 135 clclqcpdsstlvpprspgaletalpapqap-rlllll---lpvallmstawclhw 190
 QY 102 RRCRRERSPP 112
 Db 191 rrrrrrrsyp 201

Search completed: March 12, 2002, 09:56:32
 Job time: 89 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2002, 09:56:33 ; Search time 22.28 Seconds
(without alignments)
748.431 Million cell updates/sec

Title: US-09-690-454-59
Perfect score: 608
Sequence: 1 MARGSLRLRLLLVLGLWLA.....LSGLVWRRCRRRSSPPPX 114

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 473505 seqs, 146272329 residues
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_17:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phage:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	571	93.9	129	4 Q9NP84	Q9np84 homo sapien
2	450	74.0	129	11 Q9Q2W3	Q9q2w3 mus musculus
3	445	73.2	129	11 Q9CR75	Q9cr75 mus musculus
4	336.5	55.3	94	4 Q9HCS0	Q9hcs0 homo sapien
5	87	14.3	333	4 Q9BZG3	Q9bzg3 homo sapien
6	87	14.3	426	4 Q9BZG2	Q9bzg2 homo sapien
7	85.5	14.1	277	4 Q14866	Q14866 homo sapien
8	80.5	13.2	426	4 Q14865	Q14865 homo sapien
9	79	13.0	341	6 Q9NOB3	Q9nob3 macaca fasc
10	78.5	12.9	332	10 Q42839	Q42839 hordeum vul
11	75.5	12.4	222	12 Q9ORE6	Q9ore6 indian citr
12	75.5	12.4	436	2 Q53668	Q53668 mycobacteri
13	75.5	12.4	635	5 Q18288	Q18288 caenorhabdi
14	75	12.3	348	4 Q9H460	Q9h460 homo sapien
15	74.5	12.3	519	2 Q9A9X4	Q9a9x4 caulobacter
16	74	12.2	768	5 P90890	P90890 caenorhabdi
17	73.5	12.1	730	4 Q9P2P7	Q9p2p7 homo sapien
18	73	12.0	227	10 Q04393	Q04393 hordeum vul
19	73	12.0	768	13 Q98TH8	Q98th8 cyprinus ca

20	72.5	11.9	329	4 Q9NQD2	Q9nqd2 homo sapien
21	72.5	11.9	346	2 Q9RSC0	Q9rsc0 deinococcus
22	72.5	11.9	365	4 Q9Y288	Q9y288 homo sapien
23	72.5	11.9	370	4 Q9BZW8	Q9bzw8 homo sapien
24	72	11.8	368	13 Q9IAR7	Q9iar7 gallus gall
25	71.5	11.8	210	4 Q9S054	Q9s054 homo sapien
26	71.5	11.8	2873	12 Q93072	Q93072 hepatitis g
27	71	11.7	116	12 Q90631	Q90631 baboon herp
28	71	11.7	175	2 Q9S255	Q9s255 streptomyce
29	71	11.7	557	11 Q9Q2F2	Q9q2f2 mus musculu
30	70.5	11.6	412	2 Q9FCA1	Q9fca1 streptomyce
31	70	11.5	206	12 Q75637	Q75637 human immun
32	70	11.5	293	6 Q9TTI0	Q9tti0 sus scrofa
33	70	11.5	301	4 Q9UJA2	Q9uja2 homo sapien
34	70	11.5	443	11 Q9JMO6	Q9jmo6 mus musculu
35	69.5	11.4	115	12 Q65536	Q65536 baboon herp
36	69.5	11.4	206	12 Q75638	Q75638 human immun
37	69.5	11.4	424	2 Q9RSY4	Q9rsy4 deinococcus
38	69.5	11.4	587	2 Q53417	Q53417 mycobacteri
39	69.5	11.4	829	5 P90662	P90662 aedes aegyp
40	69.5	11.4	1223	5 Q9VRC9	Q9vrc9 drosophila
41	69	11.3	206	12 Q75624	Q75624 human immun
42	69	11.3	401	5 Q97407	Q97407 anopheles g
43	69	11.3	539	10 Q9LGA7	Q9lg47 oryza sativ
44	69	11.3	1276	6 Q9GK52	Q9gk52 canis famil
45	68.5	11.3	484	1 Q9V083	Q9v083 pyrococcus

ALIGNMENTS

RESULT 1
Q9NP84 ID Q9NP84 PRELIMINARY; PRT; 129 AA.
AC Q9NP84;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE TYPE I TRANSMEMBRANE PROTEIN PRECURSOR (TYPE I TRANSMEMBRANE PROTEIN FN14).
GN FN14.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tanaka S., Sugimachi K.;
RT "Human homologue of Fn14.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=20216634; PubMed=10751351;
RX TISSUE=PLACENTA;
RA Feng S.-L.Y., Guo Y., Factor V.M., Thorgeirsson S.S., Bell D.W.,
RA Testa J.R., Peiffley K.A., Winkles J.A.;
RT "The Fn14 immediate-early response gene is induced during liver
regeneration and highly expressed in both human and murine
hepatocellular carcinoma.";
RL Am. J. Pathol. 156:1253-1261(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=ENDOMETRIAL ADENOCARCINOMA;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB035480; BAA94792.1; -;
DR EMBL; AF1911148; AAF69108.1; -;
DR EMBL; BC002718; AAH02718.1; -;
KW Signal; Transmembrane.
FT SIGNAL 9 31 POTENTIAL.
SQ SEQUENCE 129 AA; 13911 MW; BF3FDFB9C1E1C448 CRC64;

RESULT	7
ID	O14866 PRELIMINARY; PRT; 277 AA.
AC	O14866;
DT	01-JAN-1998 (TrEMBLrel. 05, Created)
DT	01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE	SOLUBLE DEATH RECEPTOR 3 BETA.
GN	DR3.
GN	Human sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
[1]	
RN	SEQUENCE FROM N.A.
RP	Best Local Similarity
RA	Walzocha K., Ribeiro P., Renard N., Charlot C., Coiffier B.,
RA	Salles G.;
RL	Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF026071; AAB82288.1; -
DR	HSSP; P19438; INCF.
DR	InterPro; IPR000561; EGF-like.
DR	InterPro; IPR001368; TNFR_c6.
DR	Pfam; PF00020; TNFR_c6; 2.
DR	SMART; SM00208; TNFR; 2.
DR	PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR	PROSITE; PS00652; TNFR_NGFR_1; 2.
DR	PROSITE; PS00050; TNFR_NGFR_2; 1.
SQ	SEQUENCE 277 AA; 29111 MW; 3D19F3E847BFC093 CRC64;
Query Match	14.1%; Score 85.5; DB 4; Length 277;
Best Local Similarity	26.2%; Pred. No. 0.45; Mismatches 11; Indels 43; Gaps
Matches	32; Conservative 11;
QY	4 GSLRRLRLVLGLWLLRSVAGEAPGTAPCGSGSSADLDKCMDCASCRA--RPHS 61 : : : : : : : : : : : : Db 145 GALRHTRL-----CSRRT-----DCGTCLPGFYHG 173
QY	62 DFLCAGCAAPPAPRLLWPILGGALSITF-----VLGLLSGLFW---RRCRERSSP 111 : : : : : : : : : : : : Db 174 DGCVCSTPTPPSLAGAPGWAVQSAPLVSVAGRGVGLGMRVGELGWTEGRVRRCATTQ 233
QY	112 PP 113
Db	234 HP 235
RESULT	8
ID	O14865 PRELIMINARY; PRT; 426 AA.
AC	O14865;
DT	01-JAN-1998 (TrEMBLrel. 05, Created)
DT	01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE	DEATH RECEPTOR 3 BETA.
GN	DR3.
GN	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
[1]	
RN	SEQUENCE FROM N.A.
RP	MEDLINE=98113360; PubMed=9446802;
RX	Walzocha K., Ribeiro P., Charlot C., Renard N., Coiffier B.,
RA	Salles G.;
RT	"A new death receptor 3 isoform: expression in human lymphoid cell
RT	lines and non-Hodgkin's lymphomas";
RL	Biochem. Biophys. Res. Commun. 242:376-379(1998).
DR	EMBL; AF026070; AAC39556.1; -
DR	HSSP; P19438; INCF.
DR	InterPro; IPR000488; Death.
DR	InterPro; IPR000561; EGF-like.

DR InterPro: IPR001368; TNFR_c6.
DR Pfam: PF00531; death; 1.
DR Pfam: PF00020; TNFR_c6; 2.
DR SMART: SM00005; DEATH; 1.
DR SMART: SM00208; TNFR; 2.
DR PROSITE: PS0017; DEATH_DOMAIN; 1.
DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
DR PROSITE: PS00652; TNFR_NGFR_1; 2.
DR PROSITE: PS00050; TNFR_NGFR_2; 1.
SQ SEQUENCE 426 AA; 45950 MW; 371AA7F16AD29C16 CRC64;

Query Match 13.2%; Score 80.5; DB 4; Length 426;

Best Local Similarity 22.2%; Pred. No. 2.1;
Matches 32; Conservative 13; Mismatches 34; Indels 65; Gaps 6;

Qy 4 GSLRLRLVLLGLWLLALLRSVAGEQAPCTAPCSRSGSSWSADLDKCMDCASCRA--RPHS 61
Db 145 GALHRRTRL-----CSRSDT-----DCGTCLPGFYEHG 173
Qy 62 DFCLGCAAPAPFRLW-----PILGGALSLTFVLGLSGFLV-----W 101
Db 174 DGVSCPTPPPSLAGPWGAVQSAPLSVAGRGVGVWQVLLAGLVVLLGLGATLTYY 233

Qy 102 RRC-----RRRSPPP 113
Db 234 RHCWPKPLVTADAGMEALTPP 257

RESULT 9

ID Q9NOB3 PRELIMINARY; PRT; 341 AA.
AC Q9NOB3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE GAP JUNCTION PROTEIN (CONNEXIN).
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
libraries."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ONE GAP JUNCTION CONSISTS OF A CLUSTER OF CLOSELY PACKED
CC PAIRS OF TRANSMEMBRANE CHANNELS. THE CONNEXONS, THROUGH WHICH
CC MATERIALS OF LOW MW DIFFUSE FROM ONE CELL TO A NEIGHBOURING CELL
CC (BY SIMILARITY).
CC -!- SUBUNIT: A CONNEXON IS COMPOSED OF AN HEXAMER OF CONNEXINS (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: TO THE EUKARYOTIC CONNEXIN FAMILY.
DR EMBL: AB046017; BAB01599.1; -.
DR InterPro: IPR000500; Connexin.
DR Pfam: PF00029; connexin; 1.
DR PRINTS: PR00206; CONNEXIN.
DR SMART: SM00037; CNX; 1.
DR PROSITE: PS00407; CONNEXINS_1; 1.
DR PROSITE: PS00408; CONNEXINS_2; 1.
SQ Gap junction; Transmembrane.
KW SEQUENCE 341 AA; 36784 MW; 4A71DEA938F0FA0C CRC64;

Query Match 13.0%; Score 79; DB 6; Length 341;

Best Local Similarity 27.9%; Pred. No. 2.4;
Matches 29; Conservative 9; Mismatches 46; Indels 20; Gaps 1;

Qy 9 LRLVLLGLWLLALLRSVAGEQAPCTAPCSRSGSSWSADLDKCMDCASCRAAPHSDFLGCA 68

Db 119 LLRTLLEAAGALNYLLFGFLAPNKFPT-----RPCTGVDCY 158
Qy 69 AAPAPERLLWPIILGGALSITFVLGLSGFLVWRRRCRRERSPP 112
Db 159 VSRTEKSLMLFLWAVSALSFLGLGLADLVCSLRLMRPP 202

RESULT 10

Q42839
ID Q42839 PRELIMINARY; PRT; 332 AA.
AC Q42839;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CHITINASE (EC 3.2.1.14) (CHITODEXTRINASE) (1,4-BETA-POLY-N-
DE ACETYLGLUCOSAMINIDASE) (POLY-BETA-GLUCOSAMINIDASE).
GN CHIAL OR CH133.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SEED;
RX MEDLINE=95078949; PubMed=7987416;
RA Leah R., Skriver K., Knudsen S., Ruud-Hansen J., Raikhel N.V.,
RA Mundy J.;
RT "Identification of an enhancer/silencer sequence directing the
RT aleurone-specific expression of a barley chitinase gene."
RL Plant J. 6:579-589(1994).
CC -!- CATALYTIC ACTIVITY: HYDROLYSIS OF THE 1,4-BETA-LINKAGES OF N-
CC ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.
DR EMBL: L34211; AA56787.1; -.
DR HSP: P23951; 2BAA.
DR Mendel; 1565; Horvu; Chial; 1565.
DR InterPro: IPR001002; Chitin_bind.
DR InterPro: IPR000726; Glyco_hydro_19.
DR Pfam: PF00187; chitin_binding; 1.
DR Pfam: PF00182; Glyco_hydro_19; 1.
DR PRINTS: PR00451; CHITINBINDNG.
DR ProDom: PD000574; Glyco_hydro_19; 1.
DR ProDom: PD000609; Chitin_bind; 1.
DR SMART: SM00270; ChitBD1; 1.
DR PROSITE: PS00773; CHITINASE_19_1; UNKNOWN_1.
DR PROSITE: PS00774; CHITINASE_19_2; 1.
DR PROSITE: PS00026; CHITIN_BINDING; 1.
KW Chitin-binding; Glycosidase; Hydrolase.
SQ SEQUENCE 332 AA; 35369 MW; 1B38554F49AC9E0F CRC64;

Query Match 12.9%; Score 78.5; DB 10; Length 332;
Best Local Similarity 32.0%; Pred. No. 2.7;
Matches 33; Conservative 9; Mismatches 38; Indels 23; Gaps 7;

Qy 13 LVLLGLWLLALLRSV-----AGEQAPG-TAP-----CSRSGSSWSADLDKCMDCASCRAPHSDF 63
Db 13 IVLSAALAMAMVVRQAQCGSQAGGATCPNCCSRFGYCGSTSYC--GAGCQSQ----- 65

Qy 64 CLGCAAPAPFRLWPIILGGALSITFVLGLSGFLVWRR-RCR 105

Db 66 CSGCGPTPPGSP-----GGGVSIISRDLFEQFLHHRDRCQ 102

RESULT 11

Q9QEE6
ID Q9QEE6 PRELIMINARY; PRT; 222 AA.
AC Q9QEE6;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 23 KDA PROTEIN.

OS Indian citrus ringspot virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage.
OX NCBI_TaxID=104664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RA Rustici G., Accotto G.P., Noris E., Masenga V., Luisoni E.,
RA Milne R.G.;
RT "Indian citrus ringspot virus: a proposed new species with some
RT affinities to potex-, carla-, fovea- and allexiviruses.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF184962; AAF01314.1; -
SQ SEQUENCE 222 AA; 25455 MW; 636A47058DD8D01D CRC64;

Query Match 12.4%; Score 75.5; DB 12; Length 222;
Best Local Similarity 35.4%; Pred. No. 3.8;
Matches 35; Conservative 7; Mismatches 38; Indels 19; Gaps 7;

QY 3 RGSRLRLRLVLGLLALLRSVAGEQAP-CTAPCSRGSWSADLDKCMDCASCRA--RP 59
Db 21 RGSQIRSVRLP---WRPFRFPV---CPSGTSPYSRGTHSQPSYVRQCNCERARQWFR 74
QY 60 H-SDFCL-----GCAAAPAPFRL---WPILGGALS 88
Db 75 HDGPRCLHQRPDYSRLQAPDPDPFQHLNSFEFILLALS 113

RESULT 12
OS3668 PRELIMINARY; PRT; 436 AA.
AC O53668;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DE PUTATIVE MEMBRANE PROTEIN.
GN RV0246 OR MTV034.12.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
RA Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
DR EMBL; AL021929; CAAL1738.1; -
DR TuberculList; Rv0246; -
KW Complete proteome.
SQ SEQUENCE 436 AA; 45325 MW; 393ABC2532F42D2A CRC64;

Query Match 12.4%; Score 75.5; DB 2; Length 436;
Best Local Similarity 23.0%; Pred. No. 6.7;
Matches 35; Conservative 16; Mismatches 52; Indels 49; Gaps 5;

QY 1 MARGSLRLRLRLVLGLWALLRSVAGEQA--POTAPCSRGS--WSADLDKCMDCASCRC 56
Db 142 MPGSTARARILLTEVGVGAATAVVAATLSFVPDQHLPSRLHLLWTAAVAMATSAACR 201
QY 57 ARPHSDFCLGCAAPAPFLLW-----PILGG----- 84
Db 202 ALPHR--IVPRVHAAPGLHKLVLVGVGTARTNGWRYRYLLVQVLFGSVVLGSSFSIRVA 259

QY 85 -----ALSLTFVLGLLSGFLVWRRCR 105
Db 260 AVPGDQPDVAVVLFVVCVGLLGGIALNVRV 291

RESULT 13
O18288 PRELIMINARY; PRT; 635 AA.
AC O18288;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE ZK1010.9 PROTEIN.
GN ZK1010.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Gardner A., McMurray A.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z81503; CAB04115.1; -
DR EMBL; Z82083; CAB04115.1; JOINED.
DR EMBL; Z82083; CAB04975.1; -
DR EMBL; Z81503; CAB04975.1; JOINED.
DR InterPro; IPR00175; Na_neurotran_sympot.
DR Pfam; PF00209; SNF; 2.
DR PRINTS; PR00176; NANEUSMPORT.
DR PROSITE; PS0267; NA_NEUROTRAN_SYMP_3; 1.
KW Hypothetical protein.
SQ SEQUENCE 635 AA; 71267 MW; 2841834039D8F14A CRC64;

Query Match 12.4%; Score 75.5; DB 5; Length 635;
Best Local Similarity 36.6%; Pred. No. 9.2;
Matches 26; Conservative 10; Mismatches 28; Indels 7; Gaps 3;

QY 31 PGTAPCSRSWSADLDKCMDCASCARP--HSDFCLG--CAAAPAPFRLWLPILGGAL 86
Db 92 PTAYKNGLSF---LIAVVCGLLFAVPAIHMEFALGQYAKSPPAERFMMPILEGVG 148
QY 87 SLTFVLGLLSG 97
Db 149 WMTCLVGAIG 159

RESULT 14
O9H460 PRELIMINARY; PRT; 348 AA.
AC O9H460;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE BA425A6.2 (SIMILAR TO CONNEXIN) (FRAGMENT).
GN BA425A6.2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Heath P.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL121749; CAC10186.1; -
DR InterPro; IPR000500; Connexin.
DR Pfam; PF00029; connexin; 1.
DR PRINTS; PR00206; CONNEXIN.
DR SMART; SM00037; CNX; 1.
DR PROSITE; PS00407; CONNEXINS_1; 1.
FT NON_TER 1
SQ SEQUENCE 348 AA; 37805 MW; 07EB6478067F8995 CRC64;

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	72	11.8	3724	2	US-08-804-227C-10	Sequence 10, Appl	
2	72	11.8	3724	2	US-08-804-186-4	Sequence 4, Appli	
3	69	11.3	215	4	US-09-220-528-104	Sequence 104, App	
4	68.5	11.3	4205	2	US-08-775-009-37	Sequence 37, Appl	
5	68.5	11.3	429	2	US-07-964-589-2	Sequence 2, Appli	
6	68.5	11.3	429	5	PCT-US93-02024-2	Sequence 2, Appli	
7	67.5	11.1	174	4	US-09-383-586-12	Sequence 12, Appl	
8	67.5	11.1	359	4	US-09-413-814-90	Sequence 90, Appl	
9	67.5	11.1	361	4	US-09-413-814-77	Sequence 77, Appl	
10	67	11.0	928	1	US-08-442-248-2	Sequence 2, Appli	
11	67	11.0	928	1	US-08-440-815-2	Sequence 2, Appli	
12	67	11.0	928	4	US-08-485-449-2	Sequence 2, Appli	
13	67	11.0	1005	2	US-08-469-537A-103	Sequence 103, App	
14	67	11.0	1481	2	US-08-616-844-40	Sequence 40, Appl	
15	67	11.0	1481	2	US-08-599-654-40	Sequence 40, Appl	
16	67	11.0	1481	3	US-08-944-868A-40	Sequence 40, Appl	
17	67	11.0	1481	3	US-08-944-423A-40	Sequence 40, Appl	
18	67	11.0	1481	3	US-08-944-496-40	Sequence 40, Appl	
19	66	10.9	368	2	US-08-651-579-2	Sequence 2, Appli	
20	66	10.9	992	1	US-08-127-499A-1	Sequence 1, Appli	
21	66	10.9	992	1	US-08-482-847-1	Sequence 1, Appli	
22	66	10.9	1238	4	US-09-214-278-5	Sequence 5, Appli	
23	65.5	10.8	2910	1	US-08-466-033-183	Sequence 183, App	
24	65.5	10.8	2910	2	US-08-444-733-183	Sequence 183, App	
25	65.5	10.8	2910	2	US-08-464-134-183	Sequence 183, App	
26	65.5	10.8	2910	2	US-08-461-361-183	Sequence 183, App	
27	65.5	10.8	2910	2	US-08-485-910-183	Sequence 183, App	

; Sequence 12, Application US/09383586
; Patent No. 6242419
; GENERAL INFORMATION:
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Abernethy, Nevin
; APPLICANT: Onrust, Rene
; APPLICANT: Kumble, Anand
; APPLICANT: Murison, Greg
; TITLE OF INVENTION: Compounds isolated from stromal cells
; TITLE OF INVENTION: and methods for their use
; FILE REFERENCE: 11000.1037c1
; CURRENT APPLICATION NUMBER: US/09/383,586
; CURRENT FILING DATE: 1999-08-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Mouse
US-09-383-586-12

Query Match 11.1%; Score 67.5; DB 4; Length 174;
Best Local Similarity 25.4%; Pred. No. 3.9;
Matches 30; Conservative 14; Mismatches 53; Indels 21; Gaps 4;
Qy 12 LVLGLWLLALLRSVAGEQAPGTAPCSR-----WSADLCKMDCASCRR 58
Db 13 VMLGAVLLLLLSGASQAPPRVGCSEYTNRSCELRNVSLWCNENKACMDYPVRKIL 72
Qy 59 PHSDFCGCAAAAPPAPFRLWLPILGALSUTFVLG--LLSGFLV--WRRCRRRRSPP 112
Db 73 PPASLC-----KLSSRWGVCWVNFALITMSVLGSGVLLGITVCCYCCRRKKSRP 126

RESULT 8
US-09-413-814-90
; Sequence 90, Application US/09413814
; Patent No. 6225064
; GENERAL INFORMATION:
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
; APPLICANT: Bristol-Myers Squibb, Co.
; APPLICANT: Beyer, Stefan
; APPLICANT: Bloecker, Helmut
; APPLICANT: Brandt, Petra
; APPLICANT: Cino, Paul M
; APPLICANT: Dougherty, Brian A
; APPLICANT: Goldberg, Steven L
; APPLICANT: Hofle, Gerhard
; APPLICANT: Mueller, Joachim
; APPLICANT: Reichenbach, Hans
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
; TITLE OF INVENTION: heteropolyketide compounds
; FILE REFERENCE: PCT/US 99/23535
; CURRENT APPLICATION NUMBER: US/09/413,814
; CURRENT FILING DATE: 1999-10-07
; EARLIER APPLICATION NUMBER: DE 198 46 493.2
; EARLIER FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 90
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-413-814-90

Query Match 11.1%; Score 67.5; DB 4; Length 359;
Best Local Similarity 24.3%; Pred. No. 8.9;
Matches 25; Conservative 12; Mismatches 31; Indels 35; Gaps 4;
Qy 11 RLLVLGLWLLALLRSVAGEQAP-----GTAPCSR-----GSSWS 43

Db 243 RIVLLALWALLMRMSGNEPEVTVAVRFDGRSLDALAGAVGPFARFLPVRIEISADTL 302
Qy 44 ADLDKCMDCASCRRPHSDFCGCAAAAPPAPFRLWLPIL--GG 84
Db 303 ADVARRLALAEAAAHQD-----AAPGVSHRMSWGLLRGG 339
RESULT 9
US-09-413-814-77
; Sequence 77, Application US/09413814
; Patent No. 6225064
; GENERAL INFORMATION:
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
; APPLICANT: Bristol-Myers Squibb, Co.
; APPLICANT: Beyer, Stefan
; APPLICANT: Bloecker, Helmut
; APPLICANT: Brandt, Petra
; APPLICANT: Cino, Paul M
; APPLICANT: Dougherty, Brian A
; APPLICANT: Goldberg, Steven L
; APPLICANT: Hofle, Gerhard
; APPLICANT: Mueller, Joachim
; APPLICANT: Reichenbach, Hans
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
; TITLE OF INVENTION: heteropolyketide compounds
; FILE REFERENCE: PCT/US 99/23535
; CURRENT APPLICATION NUMBER: US/09/413,814
; CURRENT FILING DATE: 1999-10-07
; EARLIER APPLICATION NUMBER: DE 198 46 493.2
; EARLIER FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 77
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-413-814-77

Query Match 11.1%; Score 67.5; DB 4; Length 361;
Best Local Similarity 24.3%; Pred. No. 8.9;
Matches 25; Conservative 12; Mismatches 31; Indels 35; Gaps 4;
Qy 11 RLLVLGLWLLALLRSVAGEQAP-----GTAPCSR-----GSSWS 43
Db 245 RIVLLALWALLMRMSGNEPEVTVAVRFDGRSLDALAGAVGPFARFLPVRIEISADTL 304
Qy 44 ADLDKCMDCASCRRPHSDFCGCAAAAPPAPFRLWLPIL--GG 84
Db 305 ADVARRLALAEAAAHQD-----AAPGVSHRMSWGLLRGG 341
RESULT 10
US-08-442-248-2
; Sequence 2, Application US/08442248
; Patent No. 5759863
; GENERAL INFORMATION:
; APPLICANT: Caras, Ingrid W.
; APPLICANT: Winslow, John W.
; TITLE OF INVENTION: AL-1 Neurotrophic Factor
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/442,248
FILING DATE: 15-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/330128
FILING DATE: 27-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: 920C4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-8674
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 928 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-442-248-2

Query Match 11.0%; Score 67; DB 1; Length 928;
Best Local Similarity 31.6%; Pred. No. 30;
Matches 18; Conservative 1; Mismatches 20; Indels 18; Gaps 3;

QY 31 PGTAPCSRGSSWSADLDKCM-----DCASCR-----ARPHSDFCLGCAAAP 72
DB 273 PPKMHCSAEGEWLVPICMKCAGYEKNGTCQVCRPGFFKASPHSQTSKC---PP 326

RESULT 11
US-08-440-815-2
Sequence 2, Application US/08440815
Patent No. 5798448
GENERAL INFORMATION:
APPLICANT: Caras, Ingrid W.
APPLICANT: Winslow, John W.
TITLE OF INVENTION: AL-1 Neurotrophic Factor
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,815
FILING DATE: 15-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/330128
FILING DATE: 27-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: 920C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-8674
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 928 amino acids
TYPE: amino acid

TOPOLOGY: linear
US-08-440-815-2

Query Match 11.0%; Score 67; DB 1; Length 928;
Best Local Similarity 31.6%; Pred. No. 30;
Matches 18; Conservative 1; Mismatches 20; Indels 18; Gaps 3;

QY 31 PGTAPCSRGSSWSADLDKCM-----DCASCR-----ARPHSDFCLGCAAAP 72
DB 273 PPKMHCSAEGEWLVPICMKCAGYEKNGTCQVCRPGFFKASPHSQTSKC---PP 326

RESULT 12
US-08-486-449-2
Sequence 2, Application US/08486449
Patent No. 6280732
GENERAL INFORMATION:
APPLICANT: Caras, Ingrid W.
APPLICANT: Winslow, John W.
TITLE OF INVENTION: AL-1 Neurotrophic Factor
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,449
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/330128
FILING DATE: 27-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: P0920P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-8674
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 928 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-486-449-2

Query Match 11.0%; Score 67; DB 4; Length 928;
Best Local Similarity 31.6%; Pred. No. 30;
Matches 18; Conservative 1; Mismatches 20; Indels 18; Gaps 3;

QY 31 PGTAPCSRGSSWSADLDKCM-----DCASCR-----ARPHSDFCLGCAAAP 72
DB 273 PPKMHCSAEGEWLVPICMKCAGYEKNGTCQVCRPGFFKASPHSQTSKC---PP 326

RESULT 13
US-08-469-537A-103
Sequence 103, Application US/08469537A
Patent No. 5843749
GENERAL INFORMATION:
APPLICANT: Maisonnier, et al.
TITLE OF INVENTION: EHK AND ROR TYROSINE

```
; TITLE OF INVENTION: KINASES
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill River Road
; CITY: Tarrytown
; STATE: NY
; COUNTRY: U.S.A.
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,537A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/406,247
; FILING DATE: 17-MAR-1995
; APPLICATION NUMBER: USSN 08/144,992
; FILING DATE: 28-OCT-1993
; APPLICATION NUMBER: USSN 07/736,559
; FILING DATE: 26-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Kempler, Ph.D., Gail M
; REGISTRATION NUMBER: 32,143
; REFERENCE/DOCKET NUMBER: REG 070C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 914-345-7400
; TELEFAX: 914-345-7721
; TELEX:
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1005 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-469-537A-103

Query Match 11.0%; Score 67; DB 2; Length 1005;
Best Local Similarity 31.6%; Pred. No. 33;
Matches 18; Conservative 1; Mismatches 20; Indels 18; Gaps 3;

QY 31 PCTAPCSRGSSWSADLDKCM-----DCASCR-----ARPHSDFCLGCAAAP 72
Db 273 PPKMHCSAEGEWLPVIGKCMKAGYEKNGTCQVCRPGFFKASPHSQTCSKC---PP 326

RESULT 14
US-08-616-844-40
; Sequence 40, Application US/08616844
; Patent No. 5849578
; GENERAL INFORMATION:
; APPLICANT: FALB, DEAN A.
; TITLE OF INVENTION: COMPOSITION AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/599,654
; FILING DATE: 09-FEB-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:

; TITLE OF INVENTION: KINASES
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill River Road
; CITY: Tarrytown
; STATE: NY
; COUNTRY: U.S.A.
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,537A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/406,247
; FILING DATE: 17-MAR-1995
; APPLICATION NUMBER: USSN 08/144,992
; FILING DATE: 28-OCT-1993
; APPLICATION NUMBER: USSN 07/736,559
; FILING DATE: 26-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Kempler, Ph.D., Gail M
; REGISTRATION NUMBER: 32,143
; REFERENCE/DOCKET NUMBER: REG 070C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 914-345-7400
; TELEFAX: 914-345-7721
; TELEX:
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1005 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-469-537A-103

Query Match 11.0%; Score 67; DB 2; Length 1481;
Best Local Similarity 24.1%; Pred. No. 51;
Matches 32; Conservative 18; Mismatches 37; Indels 46; Gaps 8;

QY 4 GSURRLRLVGLWLALRSVAGEQAPGTAP-CSRGSSWSADLDKCMDC----- 53
Db 1265 GSQRIRF-----AGSLCKRKSPECDKTSICTDLGDVALCOCKSGYFOFN 1310

QY 54 -----SCRA-----RPHSDFCLGCAAAPA-----PFRLLWPIL---GGALSITFVLGLLS 96
Db 1311 KMDHSCRACEDGYLENETCMSCFPGLGLNCGNPNYQLTIVVIAAAGGGLL--ILGIA- 1367

QY 97 GFLVWRRCRRERS 109
Db 1368 --LIVTCCRKNK 1378

RESULT 15
US-08-599-654-40
; Sequence 40, Application US/08599654
; Patent No. 5882925
; GENERAL INFORMATION:
; APPLICANT: FALB, DEAN A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/599,654
; FILING DATE: 09-FEB-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
```

/ APPLICATION NUMBER: US 08/485,573
/ FILING DATE: 07-JUN-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/386,844
/ FILING DATE: 10-FEB-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: CORUZZI, LAURA A
/ REGISTRATION NUMBER: 30,742
/ REFERENCE/DOCKET NUMBER: 7853-041
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 790-9090
/ TELEFAX: (212) 869-8864
/ TELEX: 66141 PENNIE
/ INFORMATION FOR SEQ ID NO: 40:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1481 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: unknown
/ MOLECULE TYPE: protein
/ US-08-599-654-40

Query Match 11.0%; Score 67; DB 2; Length 1481;
Best Local Similarity 24.1%; Pred. No. 51;
Matches 32; Conservative 18; Mismatches 37; Indels 46; Gaps 8;
QY 4 GSRLRLRLVLGLMLALLRSVAGEQAPGTAP-CSRGSWSADLDKCMDC----- 53
Db 1265 GSQRIRF-----AGSLCKRKSPECCKDTSICTDLGVALCQCKSGYQFN 1310
QY 54 -----SCRA-----RPHSDFCGCGAAAPPA-----PERLLWPIL---GGALSITFVLGLLS 96
Db 1311 KMDHSCRACEDGYRLENETCMSCFPGLGLNCGNPNYQLITVVIAAGGGLL--ILGIA- 1367
QY 97 GFLVWRRCRRRS 109
Db 1368 --LIVTCCRKNKN 1378

Search completed: March 12, 2002, 09:57:13
Job time: 130 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2002, 09:55:38 ; Search time 96.28 seconds
(without alignments)
328.759 Million cell updates/sec

Title: US-09-690-454-59
Perfect score: 608
Sequence: 1 MARGSLRRLRLVLGLWLA.....LSGFLVRRRCRRSSPPPX 114

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues 3148936
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/2/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	607	99.8	114	1 PCT-US98-10868-60	Sequence 60, Appl
2	607	99.8	114	15 US-09-189-144-59	Sequence 59, Appl
3	571	93.9	129	17 US-09-307-140-627	Sequence 627, App
4	571	93.9	129	17 US-09-316-633-265	Sequence 265, App
5	571	93.9	129	18 US-09-445-258-16	Sequence 16, Appl
6	571	93.9	129	18 US-09-465-587-220	Sequence 220, App
7	571	93.9	129	21 US-09-742-454A-4	Sequence 4, Appli
8	571	93.9	129	22 US-09-817-076-627	Sequence 627, App
9	571	93.9	129	22 US-09-883-777-4	Sequence 4, Appli

10	571	93.9	129	24	US-60-109-869-2	Sequence 2, Appli
11	450	74.0	129	21	US-09-742-454A-5	Sequence 5, Appli
12	450	74.0	129	22	US-09-883-777-5	Sequence 5, Appli
13	433	71.2	309	21	US-09-742-454A-7	Sequence 7, Appli
14	433	71.2	309	22	US-09-883-777-7	Sequence 7, Appli
15	379.5	62.4	300	22	US-09-883-777-9	Sequence 9, Appli
16	274.5	45.1	112	1	PCT-US99-17130-136	Sequence 136, App
17	274.5	45.1	112	18	US-09-489-847-139	Sequence 139, App
18	274.5	45.1	155	1	PCT-US99-17130-277	Sequence 277, App
19	274.5	45.1	155	18	US-09-489-847-284	Sequence 284, App
20	274.5	45.1	156	1	PCT-US99-17130-225	Sequence 225, App
21	274.5	45.1	156	18	US-09-489-847-228	Sequence 228, App
22	96.5	15.9	248	16	US-09-252-991A-29249	Sequence 29249, A
23	88.5	14.6	156	24	US-60-207-214-348	Sequence 348, App
24	88.5	14.6	400	16	US-09-252-991A-26145	Sequence 26145, A
25	87.5	14.4	631	16	US-09-252-991A-20063	Sequence 20063, A
26	87	14.3	426	1	PCT-US01-08631-53707	Sequence 53707, A
27	87	14.3	426	1	PCT-US01-11797-44	Sequence 44, Appli
28	87	14.3	426	24	US-60-224-309-1	Sequence 1, Appli
29	85	14.0	731	24	US-60-213-800-278	Sequence 278, App
30	83.5	13.7	152	16	US-09-252-991A-31619	Sequence 31619, A
31	80.5	13.2	249	16	US-09-252-991A-29850	Sequence 29850, A
32	80.5	13.2	426	1	PCT-US99-26035-5	Sequence 5, Appli
33	80.5	13.2	426	16	US-09-205-018-5	Sequence 5, Appli
34	80.5	13.2	426	18	US-09-454-910-5	Sequence 5, Appli
35	80.5	13.2	508	1	PCT-US94-14643-3	Sequence 3, Appli
36	80.5	13.2	508	5	US-08-166-333-3	Sequence 3, Appli
37	80.5	13.2	508	5	US-08-173-512-3	Sequence 3, Appli
38	79	13.0	151	22	US-09-859-500-159	Sequence 159, App
39	79	13.0	194	21	US-09-758-462-1721	Sequence 1721, Ap
40	78	12.8	128	24	US-60-311-261-1393	Sequence 1393, Ap
41	77	12.7	152	1	PCT-US01-08631-56652	Sequence 56652, A
42	77	12.7	448	20	US-09-674-330A-3	Sequence 3, Appli
43	76.5	12.6	250	17	US-09-322-409-31	Sequence 31, Appli
44	76.5	12.6	250	18	US-09-451-527-31	Sequence 31, Appli
45	76.5	12.6	276	17	US-09-322-409-26	Sequence 26, Appli

ALIGNMENTS

RESULT 1
PCT-US98-10868-60
; Sequence 60, Application PC/TUS9810868
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: 32 Human Secreted Proteins
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/10868
; FILING DATE: May 27, 1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: P2006PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
PCT-US98-10868-60

Query Match 99.8%; Score 607; DB 1; Length 114;
Best Local Similarity 100.0%; Pred. No. 4.7e-51;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARGSLRLLRLVGLWLLRSVAGEQAPGTAPCSRGSWSADLDKMDCASCRARPH 60

Db 1 MARGSLRLLRLVGLWLLRSVAGEQAPGTAPCSRGSWSADLDKMDCASCRARPH 60

Qy 61 SDFCLGCAAAPAPFRLWLPILGGALSLTFVLGSLGFLVWRRCRERSPPPP 113

Db 61 SDFCLGCAAAPAPFRLWLPILGGALSLTFVLGSLGFLVWRRCRERSPPPP 113

RESULT 2

US-09-189-144-59
; Sequence 59, Application US/09189144
; GENERAL INFORMATION:
; APPLICANT: Steven M. Ruben, et al.
; TITLE OF INVENTION: 32 Human Secreted Proteins
; FILE REFERENCE: P2006P1
; CURRENT APPLICATION NUMBER: US/09/189,144
; CURRENT FILING DATE: 1998-11-10
; EARLIER APPLICATION NUMBER: PCT/US98/10868
; EARLIER FILING DATE: May 28, 1998
; EARLIER APPLICATION NUMBER: 60/044,039
; EARLIER FILING DATE: May 30, 1997
; EARLIER APPLICATION NUMBER: 60/048,093
; EARLIER FILING DATE: May 30, 1997
; EARLIER APPLICATION NUMBER: 60/048,190
; EARLIER FILING DATE: May 30, 1997
; EARLIER APPLICATION NUMBER: 60/050,935
; EARLIER FILING DATE: May 30, 1997
; EARLIER APPLICATION NUMBER: 60/048,101
; EARLIER FILING DATE: May 30, 1997
; EARLIER APPLICATION NUMBER: 60/048,356
; EARLIER FILING DATE: May 30, 1997
; EARLIER APPLICATION NUMBER: 60/056,250
; EARLIER FILING DATE: August 29, 1997
; EARLIER APPLICATION NUMBER: 60/056,296
; EARLIER FILING DATE: August 29, 1997
; EARLIER APPLICATION NUMBER: 60/056,293
; EARLIER FILING DATE: August 29, 1997
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (114)
; OTHER INFORMATION: Xaa equals stop translation
US-09-189-144-59

Query Match 99.8%; Score 607; DB 15; Length 114;
Best Local Similarity 100.0%; Pred. No. 4.7e-51;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARGSLRLLRLVGLWLLRSVAGEQAPGTAPCSRGSWSADLDKMDCASCRARPH 60

Db 1 MARGSLRLLRLVGLWLLRSVAGEQAPGTAPCSRGSWSADLDKMDCASCRARPH 60

Qy 61 SDFCLGCAAAPAPFRLWLPILGGALSLTFVLGSLGFLVWRRCRERSPPPP 113

Db 61 SDFCLGCAAAPAPFRLWLPILGGALSLTFVLGSLGFLVWRRCRERSPPPP 113

RESULT 3

US-09-307-140-627
; Sequence 627, Application US/09307140
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Nucleic Acid Molecules Derived from a
; TITLE OF INVENTION: Human Aortic Endothelium Library
; FILE REFERENCE: MLN98-14pa
; CURRENT APPLICATION NUMBER: US/09/307,140
; CURRENT FILING DATE: 1999-05-07
; EARLIER APPLICATION NUMBER: 60/084,565
; EARLIER FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 1168
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 627
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(27)
US-09-307-140-627

Query Match 93.9%; Score 571; DB 17; Length 129;
Best Local Similarity 94.7%; Pred. No. 1.7e-47;
Matches 107; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MARGSLRLLRLVGLWLLRSVAGEQAPGTAPCSRGSWSADLDKMDCASCRARPH 60

Db 1 MARGSLRLLRLVGLWLLRSVAGEQAPGTAPCSRGSWSADLDKMDCASCRARPH 60

Qy 61 SDFCLGCAAAPAPFRLWLPILGGALSLTFVLGSLGFLVWRRCRERSPPPP 113

Db 61 SDFCLGCAAAPAPFRLWLPILGGALSLTFVLGSLGFLVWRRCREREKFTTP 113

RESULT 4

US-09-316-633-265
; Sequence 265, Application US/09316633
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Nucleic Acid Molecules Derived from a
; TITLE OF INVENTION: Human Mesangial Cell Library
; FILE REFERENCE: MLN98-16pa
; CURRENT APPLICATION NUMBER: US/09/316,633
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 60/087,052
; PRIOR FILING DATE: 1998-05-27
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 265
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(27)
US-09-316-633-265

Query Match 93.9%; Score 571; DB 17; Length 129;
Best Local Similarity 94.7%; Pred. No. 1.7e-47;
Matches 107; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 MARGSLRRLRLVLGLWLLALLRSVAGEQAPGTAPCSRGSWSADLDKCMDCASCRRPH 60
 |||||
 Db 1 MARGSLRRLRLVLGLWLLALLRSVAGEQAPGTAPCSRGSWSADLDKCMDCASCRRPH 60
 |||||
 QY 61 SDFCLGCAAAPPPAPFRLLPILGGALSITFVLGSLGFLVWRRRRSSPPP 113
 |||||
 Db 61 SDFCLGCAAAPPPAPFRLLPILGGALSITFVLGSLGFLVWRRRRRREKFTTP 113
 |||||

RESULT 5
 US-09-445-258-16
 ; Sequence 16, Application US/09445258
 ; GENERAL INFORMATION:
 ; APPLICANT: Kato, Seishi
 ; APPLICANT: Sekine, Shingo
 ; APPLICANT: Kimura, Tomoko
 ; TITLE OF INVENTION: HUMAN PROTEINS HAVING TRANSMEMBRANE
 ; TITLE OF INVENTION: DOMAINS AND DNAS ENCODING THESE PROTEINS
 ; FILE REFERENCE: GIN-6706CPU5
 ; CURRENT APPLICATION NUMBER: US/09/445,258
 ; PRIOR FILING DATE: 1999-12-01
 ; PRIOR APPLICATION NUMBER: PCT/US98/02445
 ; PRIOR FILING DATE: 1998-06-03
 ; PRIOR APPLICATION NUMBER: JP 9-144948
 ; PRIOR FILING DATE: 1997-06-03
 ; NUMBER OF SEQ ID NOS: 54
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 16
 ; LENGTH: 129
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-445-258-16

Query Match 93.9%; Score 571; DB 18; Length 129;
 Best Local Similarity 94.7%; Pred. No. 1.7e-47;
 Matches 107; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 MARGSLRRLRLVLGLWLLALLRSVAGEQAPGTAPCSRGSWSADLDKCMDCASCRRPH 60
 |||||
 Db 1 MARGSLRRLRLVLGLWLLALLRSVAGEQAPGTAPCSRGSWSADLDKCMDCASCRRPH 60
 |||||
 QY 61 SDFCLGCAAAPPPAPFRLLPILGGALSITFVLGSLGFLVWRRRRSSPPP 113
 |||||
 Db 61 SDFCLGCAAAPPPAPFRLLPILGGALSITFVLGSLGFLVWRRRRRREKFTTP 113
 |||||

RESULT 6
 US-09-465-587-220
 ; Sequence 220, Application US/09465587
 ; GENERAL INFORMATION:
 ; APPLICANT: Gearing, David P.
 ; APPLICANT: Holtzman, Douglas A.
 ; APPLICANT: Robison, Keith E.
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
 ; TITLE OF INVENTION: HUMAN ADULT EPIDERMAL KERATINOCYTE LIBRARY
 ; FILE REFERENCE: 1600.1080-001
 ; CURRENT APPLICATION NUMBER: US/09/465,587
 ; CURRENT FILING DATE: 1999-12-17
 ; PRIOR APPLICATION NUMBER: 60/114,591
 ; PRIOR FILING DATE: 1998-12-31
 ; NUMBER OF SEQ ID NOS: 356
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 220
 ; LENGTH: 129
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SIGNAL
 ; LOCATION: (1)...(27)
 US-09-465-587-220

Query Match 93.9%; Score 571; DB 18; Length 129;
 Best Local Similarity 94.7%; Pred. No. 1.7e-47;
 Matches 107; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 MARGSLRRLRLVLGLWLLALLRSVAGEQAPGTAPCSRGSWSADLDKCMDCASCRRPH 60
 |||||
 Db 1 MARGSLRRLRLVLGLWLLALLRSVAGEQAPGTAPCSRGSWSADLDKCMDCASCRRPH 60
 |||||
 QY 61 SDFCLGCAAAPPPAPFRLLPILGGALSITFVLGSLGFLVWRRRRSSPPP 113
 |||||
 Db 61 SDFCLGCAAAPPPAPFRLLPILGGALSITFVLGSLGFLVWRRRRRREKFTTP 113
 |||||

RESULT 7
 US-09-742-454A-4
 ; Sequence 4, Application US/09742454A
 ; GENERAL INFORMATION:
 ; APPLICANT: WILEY, Steven R.
 ; TITLE OF INVENTION: TWEAK Receptor
 ; FILE REFERENCE: 2968-B
 ; CURRENT APPLICATION NUMBER: US/09/742,454A
 ; CURRENT FILING DATE: 2000-12-19
 ; PRIOR APPLICATION NUMBER: 60/172,878
 ; PRIOR FILING DATE: 1999-12-20
 ; PRIOR APPLICATION NUMBER: 60/203,347
 ; PRIOR FILING DATE: 2000-05-10
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 4
 ; LENGTH: 129
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-742-454A-4

Query Match 93.9%; Score 571; DB 21; Length 129;
 Best Local Similarity 94.7%; Pred. No. 1.7e-47;
 Matches 107; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 MARGSLRRLRLVLGLWLLALLRSVAGEQAPGTAPCSRGSWSADLDKCMDCASCRRPH 60
 |||||
 Db 1 MARGSLRRLRLVLGLWLLALLRSVAGEQAPGTAPCSRGSWSADLDKCMDCASCRRPH 60
 |||||
 QY 61 SDFCLGCAAAPPPAPFRLLPILGGALSITFVLGSLGFLVWRRRRSSPPP 113
 |||||
 Db 61 SDFCLGCAAAPPPAPFRLLPILGGALSITFVLGSLGFLVWRRRRRREKFTTP 113
 |||||

RESULT 8
 US-09-817-076-627
 ; Sequence 627, Application US/09817076
 ; GENERAL INFORMATION:
 ; APPLICANT: Gearing, David P.
 ; APPLICANT: Holtzman, Douglas A.
 ; APPLICANT: Robison, Keith E.
 ; TITLE OF INVENTION: Nucleic Acid Molecules Derived from a
 ; TITLE OF INVENTION: Human Aortic Endothelium Library
 ; FILE REFERENCE: MLN98-14Pa
 ; CURRENT APPLICATION NUMBER: US/09/817,076
 ; CURRENT FILING DATE: 2001-03-26
 ; PRIOR APPLICATION NUMBER: 09/307,140
 ; PRIOR FILING DATE: 1999-05-07
 ; PRIOR APPLICATION NUMBER: 60/084,565
 ; PRIOR FILING DATE: 1998-05-07
 ; NUMBER OF SEQ ID NOS: 1168
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 627
 ; LENGTH: 129
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SIGNAL
 ; LOCATION: (1)...(27)
 ;

US-09-817-076-627

Query Match 93.9%; Score 571; DB 22; Length 129;
Best Local Similarity 94.7%; Pred. No. 1.7e-47;
Matches 107; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MARGSLRLLRLVGLWLLRSVAGEQAPGTAPCSRGSWSADLDKCMDCASCARPH 60
|||||
Db 1 MARGSLRLLRLVGLWLLRSVAGEQAPGTAPCSRGSWSADLDKCMDCASCARPH 60
|||||
Qy 61 SDFCLGCAAAAPPAPFRLLPILGGALSLTFVLGSLGFLVWRRCRRSSPPP 113
|||||
Db 61 SDFCLGCAAAAPPAPFRLLPILGGALSLTFVLGSLGFLVWRRCRRREKFTTP 113
|||||

RESULT 9

US-09-883-777-4
; Sequence 4, Application US/09883777
; GENERAL INFORMATION:
; APPLICANT: Wiley, Steven R.
; TITLE OF INVENTION: TWEAK RECEPTOR
; FILE REFERENCE: 2968-C
; CURRENT APPLICATION NUMBER: US/09/883,777
; CURRENT FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/172,878
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: US 60/203,347
; PRIOR FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: PCT/US00/34755
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/742,454
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 129
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-883-777-4

Query Match 93.9%; Score 571; DB 22; Length 129;
Best Local Similarity 94.7%; Pred. No. 1.7e-47;
Matches 107; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MARGSLRLLRLVGLWLLRSVAGEQAPGTAPCSRGSWSADLDKCMDCASCARPH 60
|||||
Db 1 MARGSLRLLRLVGLWLLRSVAGEQAPGTAPCSRGSWSADLDKCMDCASCARPH 60
|||||
Qy 61 SDFCLGCAAAAPPAPFRLLPILGGALSLTFVLGSLGFLVWRRCRRSSPPP 113
|||||
Db 61 SDFCLGCAAAAPPAPFRLLPILGGALSLTFVLGSLGFLVWRRCRRREKFTTP 113
|||||

RESULT 10

US-60-109-869-2
; Sequence 2, Application US/60109869
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yue, Henry
; APPLICANT: Au-Young, Janice,
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: HUMAN MEMBRANE-SPANNING PROTEINS
; FILE REFERENCE: PF-0640 P
; CURRENT APPLICATION NUMBER: US/60/109,869
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 129

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 2454048
US-60-109-869-2

Query Match 93.9%; Score 571; DB 24; Length 129;
Best Local Similarity 94.7%; Pred. No. 1.7e-47;
Matches 107; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MARGSLRLLRLVGLWLLRSVAGEQAPGTAPCSRGSWSADLDKCMDCASCARPH 60
|||||
Db 1 MARGSLRLLRLVGLWLLRSVAGEQAPGTAPCSRGSWSADLDKCMDCASCARPH 60
|||||
Qy 61 SDFCLGCAAAAPPAPFRLLPILGGALSLTFVLGSLGFLVWRRCRRSSPPP 113
|||||
Db 61 SDFCLGCAAAAPPAPFRLLPILGGALSLTFVLGSLGFLVWRRCRRREKFTTP 113
|||||

RESULT 11

US-09-742-454A-5
; Sequence 5, Application US/09742454A
; GENERAL INFORMATION:
; APPLICANT: Wiley, Steven R.
; TITLE OF INVENTION: TWEAK RECEPTOR
; FILE REFERENCE: 2968-B
; CURRENT APPLICATION NUMBER: US/09/742,454A
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/172,878
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: 60/203,347
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-742-454A-5

Query Match 74.0%; Score 450; DB 21; Length 129;
Best Local Similarity 75.2%; Pred. No. 1.1e-35;
Matches 85; Conservative 6; Mismatches 22; Indels 0; Gaps 0;

Qy 1 MARGSLRLLRLVGLWLLRSVAGEQAPGTAPCSRGSWSADLDKCMDCASCARPH 60
|||||
Db 1 MAPGWPRLPQLVILGFLVLMRAAGEQAPGTSPCSGSSWSADLDKCMDCASCARPH 60
|||||
Qy 61 SDFCLGCAAAAPPAPFRLLPILGGALSLTFVLGSLGFLVWRRCRRSSPPP 113
|||||
Db 61 SDFCLGCAAAAPPAPFRLLPILGGALSLVLLVLSFVWRRCRRREKFTTP 113
|||||

RESULT 12

US-09-883-777-5
; Sequence 5, Application US/09883777
; GENERAL INFORMATION:
; APPLICANT: Wiley, Steven R.
; TITLE OF INVENTION: TWEAK RECEPTOR
; FILE REFERENCE: 2968-C
; CURRENT APPLICATION NUMBER: US/09/883,777
; CURRENT FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/172,878
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: US 60/203,347
; PRIOR FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: PCT/US00/34755
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/742,454
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 16

000000

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2002, 09:56:13 ; Search time 16.7 Seconds
(without alignments)
404.970 Million cell updates/sec

Title: US-09-690-454-59

Perfect score: 608

Sequence: 1 MARGSLRLLRLVLLGLWLA.....LSGFLVWRCRRRRSSPPPX 114

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283786 seqs, 5932453 residues

Total number of hits satisfying chosen parameters: 283786

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New:*

- 1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pap:*
- 2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pap:*
- 3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pap:*
- 4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pap:*
- 5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pap:*
- 6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pap:*
- 7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	607	99.8	114	5	US-09-690-454-59
2	607	99.8	114	6	US-10-062-599-59
3	87	14.3	184	6	US-10-008-063-2
4	86	14.1	133	5	US-09-708-427-65653
5	86	14.1	163	5	US-09-708-427-65652
6	84	13.8	107	5	US-09-708-427-46576
7	84	13.8	123	5	US-09-708-427-46575
8	79	13.0	719	5	US-09-708-427-65113
9	77	12.7	418	5	US-09-674-379A-3
10	75	12.3	356	1	PCT-US01-49232-74
11	75	12.3	370	1	PCT-US01-49232-73
12	73	12.0	141	5	US-09-674-266A-227
13	73	12.0	370	5	US-09-611-526-2573
14	72.5	11.9	146	5	US-09-708-427-48572
15	72.5	11.9	242	5	US-09-604-693A-254
16	72.5	11.9	242	5	US-09-605-703B-2058
17	72.5	11.9	365	6	US-10-007-303-59
18	70	11.5	89	5	US-09-708-427-80730
19	70	11.5	461	5	US-09-674-379A-8
20	70	11.5	567	5	US-09-635-949-2
21	70	11.5	567	6	US-10-004-551-2
22	70	11.5	720	5	US-09-635-949-4
23	70	11.5	720	5	US-09-989-723-231
24	70	11.5	720	5	US-09-989-724-231
25	70	11.5	720	5	US-09-989-730-231
26	70	11.5	720	5	US-09-990-436-231

27	70	11.5	720	5	US-09-990-444-231	Sequence 231, App
28	70	11.5	720	5	US-09-989-721-231	Sequence 231, App
29	70	11.5	720	5	US-09-989-722-231	Sequence 231, App
30	70	11.5	720	5	US-09-989-725-231	Sequence 231, App
31	70	11.5	720	5	US-09-989-726-231	Sequence 231, App
32	70	11.5	720	5	US-09-989-727-231	Sequence 231, App
33	70	11.5	720	5	US-09-989-728-231	Sequence 231, App
34	70	11.5	720	5	US-09-989-731-231	Sequence 231, App
35	70	11.5	720	5	US-09-989-732-231	Sequence 231, App
36	70	11.5	720	5	US-09-989-734-231	Sequence 231, App
37	70	11.5	720	5	US-09-990-437-231	Sequence 231, App
38	70	11.5	720	5	US-09-990-438-231	Sequence 231, App
39	70	11.5	720	5	US-09-990-440-231	Sequence 231, App
40	70	11.5	720	5	US-09-990-441-231	Sequence 231, App
41	70	11.5	720	5	US-09-989-279-231	Sequence 231, App
42	70	11.5	720	5	US-09-989-735-231	Sequence 231, App
43	70	11.5	720	5	US-09-990-439-231	Sequence 231, App
44	70	11.5	720	5	US-09-990-562-231	Sequence 231, App
45	70	11.5	720	5	US-09-990-711-231	Sequence 231, App

ALIGNMENTS

RESULT 1

US-09-690-454-59

; Sequence 59, Application US/09690454

; GENERAL INFORMATION:

; APPLICANT: Steven M. Ruben, et al.

; TITLE OF INVENTION: 32 Human Secreted Proteins

; FILE REFERENCE: P2006P1

; CURRENT APPLICATION NUMBER: US/09/690,454

; CURRENT FILING DATE: 2000-10-18

; PRIOR APPLICATION NUMBER: 09/189,144

; PRIOR FILING DATE: 1998-11-10

; PRIOR APPLICATION NUMBER: 60/044,039

; PRIOR FILING DATE: May 30, 1997

; PRIOR APPLICATION NUMBER: 60/048,093

; PRIOR FILING DATE: May 30, 1997

; PRIOR APPLICATION NUMBER: 60/048,190

; PRIOR FILING DATE: May 30, 1997

; PRIOR APPLICATION NUMBER: 60/050,935

; PRIOR FILING DATE: May 30, 1997

; PRIOR APPLICATION NUMBER: 60/048,101

; PRIOR FILING DATE: May 30, 1997

; PRIOR APPLICATION NUMBER: 60/048,356

; PRIOR FILING DATE: May 30, 1997

; PRIOR APPLICATION NUMBER: 60/056,250

; PRIOR FILING DATE: August 29, 1997

; PRIOR APPLICATION NUMBER: 60/056,296

; PRIOR FILING DATE: August 29, 1997

; PRIOR APPLICATION NUMBER: 60/056,293

; PRIOR FILING DATE: August 29, 1997

; NUMBER OF SEQ ID NOS: 229

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 59

; LENGTH: 114

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (114)

; OTHER INFORMATION: Xaa equals stop translation

US-09-690-454-59

Query Match

Best Local Similarity 99.8%; Score 607; DB 5; Length 114;

Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MARGSLRLLRLVLLGLWLA...LSGFLVWRCRRRRSSADLDKMDASCRRPH 60

DB 1 MARGSLRLLRLVLLGLWLA...LSGFLVWRCRRRRSSADLDKMDASCRRPH 60


```

: TITLE OF INVENTION: THEREBY
: FILE REFERENCE: 2750-1243p
: CURRENT APPLICATION NUMBER: US/09/708,427
: CURRENT FILING DATE: 2000-11-09
: NUMBER OF SEQ ID NOS: 85364
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 65652
: LENGTH: 163
: TYPE: PRT
: ORGANISM: Zea mays subsp. mays
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1..163
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: 1..163
: OTHER INFORMATION: Ceres Seq. ID 1929196
: US-09-708-427-65652

```

```

Query Match      14.1%; Score 86; DB 5; Length 163;
Best Local Similarity 35.2%; Pred. No. 0.064;
Matches 31; Conservative 10; Mismatches 27; Indels 20; Gaps 7;

Qy 29 QAPGTAPCGSRGSSWADLKCMDCASCARPHSPDFCLGCAAPAPFRLLPILGGALSL 88
      : : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 16 RSPPTTPVA--SEWS-----CVRCTLLNP-DHSDSCVACEASRPVEVDSPVVVGA --- 64

Qy 89 TFVLGLLSGFLVWRRCR--ERS-SPPP 113

Db 65 --ALALASP-----KRSRRKNERSASPPP 86

```

```

RESULT      6
US-09-708-427-46576
; Sequence 46576, Application us/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46576
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..107
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..107
; OTHER INFORMATION: Ceres Seq. ID 1852551
US-09-708-427-46576

```

Query Match	13.8%	Score 84;	DB 5;	Length 107;
Best Local Similarity	30.6%;	Pred. No.	0.069;	
Matches	37;	Conservative	11;	Mismatches 51; Indels 22; Gaps 6;
QY	1	MARGSLRRLRLVLGLWLALLRSVAGEQAPGTAPCSRGSSWSADLCKM--DCASCRA	57	
Dd		: : : :		
1 MAR	-----LGLLALAVAVLALASGVASQAPGPAPSADCASALAGMLGCLPYVOGGSTQG	54		
QY	58	RPHSDFCLGCAAA---PPAPFRLLPILTGALSILTFTVLGLSLGWLRRC-RRERSPP	112	
Dd	: : : : :			
55 KPDRECCAGVKALKUSPPSPASA-	-PFSGRTTACPSTSPAAPG-----CPRRAAKTTP	106		
QY	113	P	113	

Db 107 P 107

```

RESULT 7
US-09-708-427-46575
; Sequence 46575, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED D
; FILE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46575
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..123
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..123
; OTHER INFORMATION: Ceres Seq. ID 1852550
US-09-708-427-46575

```

	Query Match	13.8%;	Score 84;	DB 5;	Length 123;
	Best Local Similarity	30.6%;	Pred. No. 0.079;		
	Matches 37;	Conservative 11;	Mismatches 51;	Indels 22;	Gaps 6;
Qy	1	MARGSLRLLRLVLGLWLLALRRSVAGEOAPCTAPCRGSSWSADLDKCM--DCASCR	57		
Db	17	MAR-----LGLLALAVLALSGVASQAPGAPSDACASALAGLMGCLPYVOOGSTQG	70		
Qy	58	RPHSDFCLGCAAA-----PPAPFRLLPILGALSILTFVLGLLSGLFWRRRC--RRERSPP	112		
Db	71	KPDRECCAGVAAALKSPPPSPASA--PPSGRTTACPTSPAPG-----CPRAAKTPP	122		
Qy	113	P	113		
Db	123	P	123		

```

RESULT      8
US-09-708-427-65113
; Sequence 65113, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED D
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 65113
; LENGTH: 719
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..719
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..719
; OTHER INFORMATION: Ceres Seq. ID 1928247
US-09-708-427-65113

```

Query Match
13.0%; Score 79; DB 5; Length 719;

[illegible]

```

US-09-674-266A-227
; Sequence 227, Application US/09674266A
; GENERAL INFORMATION:
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKY, CHRISTIAN
; APPLICANT: DAHL, EDGAR
; APPLICANT: ROSENTHAL, ANDRE
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM PANCREAS TUMOR TISSUE
; FILE REFERENCE: ALBRE 3
; CURRENT APPLICATION NUMBER: US/09/674,266A
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 661
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 227
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-674-266A-227

```

```

Query Match          12.0%; Score 73; DB 5; Length 141;
Best Local Similarity 25.5%; Pred. No. 1.3;
Matches 26; Conservative 9; Mismatches 39; Indels 28; Gaps 4;

```

```

QY 30 APTGAPCSRSSWSADLDKMDKDCASCARPHSD-----FCLGCAAPAPPF 75
      ||||| : || : ||||| | | | | |
Db 24 SPGPSAQOQOQOQWSSARAPSRCARPSSRPPCPAVGRLASLYCCMVFPSPRPG 83
      ||||| : || : ||||| | | | | |
QY 76 RLL-----WPILG-GALSLTFVLGLLGGFLVWRRRCRRSSP 111
      | | | | | | | | | | | | | | | | | | | |
Db 84 RTWVHCTGWRLATGLWPLT-----CQVWGTPRKQOPLP 117
      | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 13
US-09-611-526-2573
; Sequence 2573, Application US/09611526
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: KAWAI, YURI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: KOJIMA, SHINICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIS OF FULL LENGTH cDNAs
; FILE REFERENCE: 08335/0122
; CURRENT APPLICATION NUMBER: US/09/611,526
; PRIOR APPLICATION NUMBER: JP 1999-194486
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: JP 2000-118774
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183765
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 484
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2573
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-611-526-2573

```

```

Query Match          12.0%; Score 73; DB 5; Length 370;
Best Local Similarity 27.9%; Pred. No. 3.3;

```

```

Matches 29; Conservative 9; Mismatches 46; Indels 20; Gaps 1;
QY 9 LURLVLGLWLLRRSVAGEQAPGTAPCSRGSSWSADLDKMDKDCASCARPHSDFCLGCA 68
      ||||| : || : ||||| | | | | |
Db 148 LLRTLLERAFGALHYFLFGFLAPKKPCT-----RPCTGVVDY 187
      ||||| : || : ||||| | | | | |
QY 69 AAPPAPFRLLPILGGALSLSLTFVLGLLGGFLVWRRRCRRSSPP 112
      : | | | | | | | | | | | | | | | | | | | |
Db 188 VSRPTEKSLMLFLNAVGALSLLGLADLVCSLRRMRRRPGPP 231
      : | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 14
US-09-708-427-48572
; Sequence 48572, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48572
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..146
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..146
; OTHER INFORMATION: Ceres Seq. ID 1855339
US-09-708-427-48572

```

```

Query Match          11.9%; Score 72.5; DB 5; Length 146;
Best Local Similarity 26.4%; Pred. No. 1.5;
Matches 23; Conservative 11; Mismatches 46; Indels 7; Gaps 2;

```

```

QY 26 AGEQAPGTAPCSRGSSWSADLDKMDKDCASCARPHSDFCLGCAAPAPPFRLLPILGGA 85
      || : | : | : | : | : | : | : | : | : | : | : |
Db 26 AGRRVPSPSSSLGPQWRRRRRRRRRAVPSPTPTARSFRRPASPT-----SSLSSG 79
      : | | : | : | : | : | : | : | : | : | : |
QY 86 LSLTFVLGLLGGFLVWRRRCRRSSPP 112
      : | | : | : | : | : | : | : | : | : | : |
Db 80 STYFTLGAAT-WMAFRRCRRRTAPP 105

```

```

RESULT 15
US-09-604-693A-254
; Sequence 254, Application US/09604693A
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN DNA REPLICATION, PROTEIN SYNTHESIS,
; FILE REFERENCE: BGI-130CP
; CURRENT APPLICATION NUMBER: US/09/604,693A
; CURRENT FILING DATE: 2000-06-27
; NUMBER OF SEQ ID NOS: 432
; SEQ ID NO 254
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-604-693A-254

```

Query Match 11.9%; Score 72.5; DB 5; Length 242;
Best Local Similarity 30.5%; Pred. No. 2.5;
Matches 18; Conservative 7; Mismatches 17; Indels 17; Gaps 4;

Qy 40 SWSADLDKMDASC-RARPHSDF-----CLGC-----AAAPPAPFRLWPILGG 84
Db 146 AGWAPSL---FDCAACGRPGPHAFHPGVGGAVCLYCRPPGSAEVPPEALHMMVLVANG 201

Search completed: March 12, 2002, 09:59:17
Job time: 184 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2002, 09:55:03 ; Search time 14.35 seconds
(without alignments)
605.150 Million cell updates/sec

Title: US-09-690-454-59
Perfect score: 608

Sequence: 1 MARGSLRLLRLVGLWLA.....LSGLVWRRCRRRSSPPPX 114

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78.5	12.9	332	2 T04484	probable chitinase
2	76.5	12.6	245	2 S78698	probable export pr
3	75.5	12.4	436	2 B70939	hypothetical prote
4	75.5	12.4	635	2 T20910	hypothetical prote
5	74	12.2	768	2 T22758	hypothetical prote
6	73.5	12.1	281	2 G69099	probable pyruvate
7	73	12.0	227	2 T04420	ribonuclease (EC 3
8	72.5	11.9	346	2 D75303	conserved hypotet
9	72	11.8	2871	2 A55624	fibrillin-1 precu
10	71	11.7	175	2 T36798	probable transcrip
11	70	11.5	558	2 I56545	glypican precursor
12	69.5	11.4	424	2 D75330	hypothetical prote
13	69.5	11.4	587	2 T08993	hypothetical prote
14	69	11.3	324	2 JC2395	Fas antigen precu
15	68.5	11.3	431	1 B69092	conserved hypotet
16	68.5	11.3	484	2 E75138	osmoregulated prol
17	68	11.2	480	2 B61213	hypothetical prote
18	68	11.2	683	2 T00872	probable protein k
19	67.5	11.1	384	1 T46966	diheme cytochrome
20	67.5	11.1	1008	1 GNVUOK	glycoprotein precu
21	67.5	11.1	1013	2 I50615	receptor-type prot
22	67	11.0	319	2 JC4390	bone marrow stroma
23	67	11.0	893	2 S51603	receptor-like tyro
24	67	11.0	898	2 S47489	receptor tyrosine
25	67	11.0	981	2 S51604	receptor-like tyro
26	67	11.0	1005	2 S49015	receptor tyrosine
27	66.5	10.9	389	2 C82987	probable MFS trans
28	66	10.9	394	2 C36942	hypothetical prote
29	66	10.9	549	2 E83085	conserved hypotet

30	66	10.9	992	1 GNVWR3	structural polypro
31	66	10.9	1053	2 T07965	reverse transcript
32	66	10.9	1766	2 A42125	trophozoite cystel
33	66	10.9	2318	2 S45306	notch 3 protein -
34	65.5	10.8	214	2 T49381	hypothetical prote
35	65.5	10.8	245	2 B36869	probable export pr
36	65.5	10.8	245	2 G85812	flagellar biosynth
37	65.5	10.8	506	2 F83545	hypothetical prote
38	65.5	10.8	884	2 T02731	serine/threonine-s
39	65.5	10.8	2321	2 S78549	notch3 protein - h
40	65	10.7	196	2 F82989	conserved hypotet
41	65	10.7	466	2 C83407	hypothetical prote
42	65	10.7	1445	1 A48148	protein-tyrosine-p
43	64.5	10.6	123	1 C44212	structural protein
44	64.5	10.6	317	2 B82577	hypothetical prote
45	64.5	10.6	335	2 T03239	probable chitinase

RESULT 1
T04484
probable chitinase (EC 3.2.1.14) - barley
C:Species: Hordeum vulgare (barley)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C:Accession: T04484
R:Leah, R.; Skriver, K.; Knudsen, S.; Ruud-Hansen, J.; Raikhel, N.V.; Mundy, J.
Plant J. 6, 579-589, 1994
A:Title: Identification of an enhancer/silencer sequence directing the aleurone-speci
A:Reference number: Z15373; MUID:95078949
A:Accession: T04484
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-332 <LEA>
A:Cross-references: EMBL:L34211; NID:g576566; PIDN:AAA56787.1; PID:g507961
C:Genetics:
A:Gene: CHI33
C:Superfamily: lectin-related plant chitinase; hevein chitin-binding domain homology;
C:Keywords: glycosidase; hydrolase; polysaccharide degradation
F:28-69/Domain: hevein chitin-binding domain homology <HCB>
F:85-322/Domain: plant chitinase homology <PCH>

Query Match	12.9%	Score	78.5	DB	2	Length	332
Best Local Similarity	32.0%	Pred. No.	2.7				
Matches	33	Conservative	9	Mismatches	38	Indels	23
						Gaps	7

QY 13 LVGLWLALLRSV----AGEQAPG-TAP---CSRGSSWSADLDKCMDCASCRCRPHSDF 63
DB 13 IVLSAALAMAMVYRAOCCGAGGATCPNCLCCSFGYCGSTSDYC--GAGCQSQ----- 65
QY 64 CLGCAAPAPFRLLPILGLGALSILTFVLGLSLGFLVWR--RCR 105
DB 66 CSQCGTTPGPSP-----GGGVSSIIISRDLEQFLLRDRQC 102

RESULT 2
S78698
probable export protein flip precursor - Salmonella typhimurium
C:Species: Salmonella typhimurium
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C:Accession: S78698
R:Ohnishi, K.; Fan, F.; Schoenhals, G.J.; Kihara, M.; Macnab, R.M.
J. Bacteriol. 179, 6092-6099, 1997
A:Title: The FlpO, FlpP, FlpQ, and FlpR proteins of Salmonella typhimurium: putative
A:Reference number: S78696; MUID:97464436
A:Accession: S78698
A:Molecule type: DNA
A:Residues: 1-245 <OHN>
A:Cross-references: EMBL:L49021; NID:gi066860; PIDN:AAB81319.1; PID:gi066863
A:Note: the sequence of residues 42-241 and the corresponding nucleic acid sequence a
C:Genetics:

A:Gene: flip
C:Function:
A:Description: may be involved in flagellar assembly; may be involved in export of flagellin
C:Superfamily: flagellum; transmembrane protein
C:Keywords: flagellum; transmembrane protein
F:1-21/Domain: signal sequence status predicted <SIG>
F:22-245/Product: probable export protein flip #status predicted <MAT>
F:45-61/Domain: transmembrane #status predicted <TM1>
F:189-105/Domain: transmembrane #status predicted <TM2>
F:189-205/Domain: transmembrane #status predicted <TM3>
F:212-228/Domain: transmembrane #status predicted <TM4>

Query Match 12.6%; Score 76.5; DB 2; Length 245;
Best Local Similarity 30.8%; Pred. No. 3.3;
Matches 33; Conservative 9; Mismatches 38; Indels 27; Gaps 6;

QY 6 LRRLRLVLGLLWALLRSVAGEAQP--TAP--CSRGSSWSADLDKCMDCASCRRP--- 59
Db 1 MRRLLFLSLAGLW--LFSPAQAQLPGLISQPLAGGQSWLSVQTLVFTTSLTFLPAIL 58
QY 60 --HSDF-----CLGCAAAAPPAPFRLLPILGGLSLTFVL 92
Db 59 LMMTSFTRIIIVFGLRLNALGTFSPAPNQV-----LLGLALFTFFI 100

RESULT 3
B70939
hypothetical protein Rv0246 - Mycobacterium tuberculosis (strain H37Rv)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: B70939
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A: Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A: Reference number: A70500; MUID: 98295987
A: Accession: B70939
A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-436 <COL>
A: Cross-references: GB:AL021929; GB:AL123456; NID:g3242291; PIDN:CAAL17338.1; PID:ei25244
A: Experimental source: strain H37Rv
C: Genetics:
A: Gene: Rv0246

Query Match 12.4%; Score 75.5; DB 2; Length 436;
Best Local Similarity 23.0%; Pred. No. 6.4;
Matches 35; Conservative 16; Mismatches 52; Indels 49; Gaps 5;

QY 1 MARGSLRRLRLVLGLWALLRSVAGEQA--PGTAPCSRGS--WSADLDKCMDCASCR 56
Db 142 MFSGTARARILLTEVGVGAALTAVAAATLSFVDPDQPLSRNILLMTAAVAMAISAICR 201
QY 57 ARPHSDFCLGCAAAPAPFRLW-----PILGG----- 84
Db 202 ALPHR--IVPRVHAPGLHKLVTGVGTATRTNGWYRRLVLQVLFSGSVLGSFSIRVA 259
QY 85 -----ALSFTFVLGLSGFLVWRRCR 105
Db 260 AVPGDQPDDEVAVLVFCVGLLGGIALNRRVR 291

RESULT 4
T20910
hypothetical protein ZK1010.9 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
A: Accession: T20910; T27646
R: McMurray, A.

submitted to the EMBL Data Library, November 1996
A:Reference number: Z19345
A:Accession: T20910
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-635 <WIL>
A:Cross-references: EMBL:Z81503; PIDN:CA04115.1; GSPDB:GN00021; CESP:ZK1010.9
A:Experimental source: clone F14F7
R: Gardner, A.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z20398
A:Accession: T27646
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-635 <WIL>
A:Cross-references: EMBL:Z82083; PIDN:CA04975.1; GSPDB:GN00021; CESP:ZK1010.9
A:Experimental source: clone ZK1010
C: Genetics:
A: Gene: CESP:ZK1010.9
A: Map position: 3
A: Introns: 71/3; 184/2; 316/3; 403/1; 444/3; 547/3; 577/3
C: Superfamily: gamma-aminobutyric acid transporter

Query Match 12.4%; Score 75.5; DB 2; Length 635;
Best Local Similarity 36.6%; Pred. No. 8.5;
Matches 26; Conservative 10; Mismatches 28; Indels 7; Gaps 3;

QY 31 PGTAPCSRGSWSADLDKCMDCASCRRP--HSDFCLG--CAAAPPAPFRLLPILGAL 86
Db 92 PTTAKYKGGISF---LIAYVVGILFAVPAIHMEFALGQYAAKSPPAFRMRMPILGVG 148
QY 87 SLTFVLGLLSG 97
Db 149 WMTCLVGAIG 159

RESULT 5
T22758
hypothetical protein F55H12.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C:Accession: T22758
R: Dobson, R.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19610
A:Accession: T22758
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-768 <WIL>
A:Cross-references: EMBL:Z81091; PIDN:CA03142.1; GSPDB:GN00019; CESP:F55H12.1
A:Experimental source: clone F55H12
C: Genetics:
A: Gene: CESP:F55H12.1
A: Map position: 1
A: Introns: 114/3; 189/1; 301/2; 426/3; 513/1; 554/3; 657/3; 687/3; 741/1
C: Superfamily: gamma-aminobutyric acid transporter

Query Match 12.2%; Score 74; DB 2; Length 768;
Best Local Similarity 40.0%; Pred. No. 14;
Matches 20; Conservative 7; Mismatches 19; Indels 4; Gaps 2;

QY 52 CASCRRP--HSDFCLG--CAAAPPAPFRLLPILGALSLTFVLGLLSG 97
Db 153 CAFVFAVPAIHMEFALGQYAAKSPPAFRMRIMPALGEGVGMWTCIVGAVIG 202

RESULT 6
G69099
probable pyruvate synthase (EC 1.2.7.1) gamma chain - Methanobacterium thermoautotrop
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 05-May-2000

C/Keywords: coenzyme A; oxidoreductase

Query Match	11.9%	Score 72.5;	DB 2;	Length 346;
Best Local Similarity	28.8%;	Pred. No. 10;		
Matches 32;	Conservative 10;	Mismatches 52;	Indels 17;	Gaps 5;

```

DD      8/  LLVTCVGLGVL--TAGSASFWV-----RWGALVAWF GAV LKI VVHLRFAGSLFVFA 139
QY      65  LGCAAA--PPAPRLLWPIIGLGSALTFLVGLLGLSGFLVWRRCRRSSPPP 113
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      139  VGTGALPHAPPLALAVSGGAALASVALGCGAMHSTRARPELHAAAPP 189

```

Ov 74 PERLL 78

Db 251 PVKAI 255

RESULTS

T04420

ribonuclease (EC 3.1.-.-) - barley

C;Species: Hordeum vulgare (barley)

C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 22-Jun-1999

C;Accession: T04420

R; Rogers, J.C.; Rogin

A: Reference number:

A: Accession: T04A20

A:Status: preliminary

A;Molecule type: DNI

A; Residues: 1-227 <

A; Cross-references:

A; Experimental sour

C;Genetics: 3E/3 07

A; ILLIONS: 33/3; 87/

C: Keywords: hydrolase; superfamily; enzyme

bioRxiv preprint doi: <https://doi.org/10.1101/2017.06.01.148010>; this version posted June 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Query Match

Best Local Similarity:

Matches 30; Con

Qy 21 11 PCVACEO

[illegible]

Db 91 LLSSLRAEW

Qy 80 PIL-----

• •

Db 148 ALASAGVAPI

RESULT. T

D75303

conserved hypothetical

C; Species: Deinococcus

C; Date: 03-Dec-1999

C;Accession: D75303

R; White, O.; Eisen,

```

RESULT 12
D75330
C:Species: Deinococcus radiodurans (strain R1)
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: D75330
R.; M.; Shen, H.O.; Venter, J.C.; Fraser, C.M.
S.: Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896
A:Accession: D75330
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-424 <WHI>
A:Cross-references: GB:AE002036; GB:AE000513; NID:g6459766; PIDN:AAF11537.1; PID:g645
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1985
A:Map position: 1

Query Match 11.4%; Score 69.5; DB 2; Length 424;
Best Local Similarity 29.7%; Pred. No. 23;
Matches 30; Conservative 11; Mismatches 35; Indels 25; Gaps 4;

QY 3 RGSRLRLRLLLVGLMLALRSVAGBQAAPTAPCSRGSSWSADLDKCMDCASCRA----- 57
|||:|||:|||:|||:|||:|||:|||:|||:|||
Db 2 RGTRITLALLLG---AGINACRNREAOQTAPPAASAAQSQAAPRKAASASPAPTATE 58
||||:|||:|||:|||:|||:|||:|||:|||

QY 58 ----RPHSDFCLGCAAA-----PPAPFRLLNPILGGALS L 88
:|||:|||:|||:|||:|||:|||:|||:|||
Db 59 PAVSAAPAADCPLPAAVTAQPAPQP-----LSRGLGL 92
:|||:|||:|||:|||:|||:|||:|||:|||

RESULT 13
C70893
hypothetical protein Rv1069c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: C70893
R.; Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
.; Connor, R.; Davies, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:98295987
A:Accession: C70893
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-587 <COL>
A:Cross-references: GB:AL021897; GB:AL123456; NID:g3256022; PIDN:CAA17185.1; PID:g289
A:Experimental source: strain H37rv
C:Genetics:
A:Gene: Rv1069c
C:Superfamily: Mycobacterium tuberculosis hypothetical protein Rv1069c

Query Match 11.4%; Score 69.5; DB 2; Length 587;
Best Local Similarity 32.5%; Pred. No. 30;
Matches 27; Conservative 2; Mismatches 17; Indels 37; Gaps 5;

QY 68 AAAP----PAPFRLLMPI-----LGGALSITFFV-----LGLLSG----- 97
|||:|||:|||:|||:|||:|||:|||:|||
Db 27 AATPQPEPQPISTWIRHYTGTGAMGLVFVWFMSFTPSLLPRGLFQGLVSGICGATG 86
||||:|||:|||:|||:|||:|||:|||:|||

QY 98 -----FLVV--RRCRERSPPP 113
|||:|||:|||:|||:|||:|||:|||
Db 87 YGLGVFAVLVRYVRSHNSPPP 109
||||:|||:|||:|||:|||:|||:|||

```


Search completed: March 12, 2002, 09:56:54
Job time: 111 sec

Query Match 11.3%; Score 69; DB 2; Length 324;
Best Local Similarity 18.2%; Pred. No. 21;
Matches 24; Conservative 18; Mismatches 34; Indels 56; Gaps 5;

QY	30	APGTACRSRGSSWSAD---	LDKCMDCA-----			-----	SCRRAPH-----	60
Db	75	APTCPCTEGEETDRKHYS	DKRCRCACFCDG	HGHEVETNCTQTNTK	CKRKNFYCNAS	134		
QY	61	-SDPCLGCAAA-----					PPAFRLMLPILGGALS	100
Db	135	LCDHCHYCTSCGLEIDLE	PCPTRTSNTKCKKQSN	YKLLWLLIIPGLAIFV	-----		FI	187

```
Qy 101 WRRCRERSPP 112
    ::|::|
Db 188 YKRYKROPQDP 199
```

RESULT 15

B69092
 conserved hypothetical protein MTH1684 - Methanobacterium thermoautotrophicum (strain De
 C:Species: Methanobacterium thermoautotrophicum
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
 C:Accession: B69092
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
 i; Olu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiawani, N.;
 ki, S.; Church, G.W.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
 A:Reference number: A69000; MUID:98037514
 A:Accession: B69092
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-431 <MTH>
 A:CROSS-references: GB:AE000926; GB:AE000666; NID:g2622806; PIDN:AAB86156.1; PID:g2622801
 A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH1684
 C:Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ1681; ferredoxin
 F:338-389/Domain: ferredoxin [24Fe-4S] homology <FER>

Query Match 11.38; Score 68.5; DB 1; Length 431;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2002, 09:56:58 ; Search time 11.68 Seconds
(without alignments)
357.859 Million cell updates/sec

Title: US-09-690-454-59
Perfect score: 608
Sequence: 1 MARGSLRLRLLLVGLWLA.....LSGFLVWRCRRERSPPPX 114

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	78	12.8	448	1 FBL5_RAT	Q9wvh8 rattus norv
2	77	12.7	448	1 FBL5_MOUSE	Q9wvh9 mus musculus
3	76.5	12.6	245	1 FLIP_SALTY	P54709 salmonella
4	73.5	12.1	81	1 PORD_METH	P56815 methanobact
5	72	11.8	1278	1 NPCL_HUMAN	O35118 homo sapien
6	72	11.8	2871	1 FBN1_MOUSE	Q61554 mus musculus
7	70	11.5	443	1 FBL4_MOUSE	Q9wvj9 mus musculus
8	70	11.5	558	1 GPC1_RAT	P35053 rattus norv
9	69.5	11.4	495	1 MLP2_DROME	Q24400 drosophila
10	69	11.3	180	1 PTTG_HUMAN	P53801 homo sapien
11	69	11.3	314	1 Y127_HUMAN	Q14140 homo sapien
12	69	11.3	324	1 FASA_RAT	Q63199 rattus norv
13	68.5	11.3	205	1 GSC1_HUMAN	O15499 homo sapien
14	68.5	11.3	431	1 YG84_METH	O27719 methanobact
15	67.5	11.1	1008	1 VGLM_UUK	P09613 ukuniemi v
16	67.5	11.1	1013	1 EPA5_CHICK	P54753 gallus gall
17	67	11.0	319	1 BST1_RAT	Q63072 rattus norv
18	67	11.0	386	1 HXAD_MOUSE	Q62424 mus musculus
19	67	11.0	1005	1 EPA5_RAT	P54757 rattus norv
20	67	11.0	2871	1 FBN1_PIG	Q9ctv36 sus scrofa
21	66.5	10.9	443	1 FBL4_CRIGR	O55058 cricetus
22	66	10.9	394	1 Y6B9_PSEAE	P42514 pseudomonas
23	66	10.9	992	1 POLS_RUBVM	Q68563 rubella vir
24	66	10.9	2318	1 NTC3_MOUSE	Q61982 mus musculus
25	65.5	10.8	245	1 FLIP_ECOLI	P33133 escherichia
26	65	10.7	1014	1 EPB6_MOUSE	O08644 mus musculus
27	65	10.7	1445	1 PTPG_HUMAN	P23470 homo sapien
28	64.5	10.6	123	1 VST1_HEVWE	Q03499 hepatitis e
29	64.5	10.6	443	1 FBL4_HUMAN	Q95967 homo sapien
30	64	10.5	1210	1 EGFR_MOUSE	Q01279 mus musculus
31	64	10.5	1227	1 B3A3_MOUSE	P16283 mus musculus
32	63.5	10.4	253	1 KLK7_HUMAN	P49862 homo sapien
33	63.5	10.4	301	1 SCOL_HUMAN	O75880 homo sapien

34	63.5	10.4	363	1 PSPB_CANFA	P17129 canis famli
35	63.5	10.4	430	1 PYRC_MYCTU	P71809 mycobacteri
36	63.5	10.4	983	1 EPA3_MOUSE	P29319 mus musculu
37	63	10.4	405	1 B3AR_CANFA	O02662 canis famli
38	63	10.4	797	1 Y900_TREPA	O83870 treponema p
39	63	10.4	1277	1 NPCL_PIG	P56941 sus scrofa
40	63	10.4	1442	1 PTPG_MOUSE	Q05909 mus musculu
41	63	10.4	2871	1 FBN1_BOVIN	P98133 bos taurus
42	62.5	10.3	129	1 SP_HUMAN	Q03403 homo sapien
43	62.5	10.3	283	1 YC00_YEAST	P25613 saccharomyc
44	62.5	10.3	543	1 TLC2_CHIMU	Q9pjp6 chlamydia m
45	62.5	10.3	547	1 Z213_HUMAN	O14771 homo sapien

ALIGNMENTS

RESULT 1

ID	FBL5_RAT	STANDARD;	PRT;	448 AA.
AC	Q9WVH8: Q9R284;			
DT	20-AUG-2001 (Rel. 40, Created)			
DT	20-AUG-2001 (Rel. 40, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	FIBULIN-5 PRECURSOR (FIBL-5) (DEVELOPMENTAL ARTERIES AND NEURAL CREST			
DE	EGF-LIKE PROTEIN) (DANCE) (EMBRYONIC VASCULAR EGF REPEAT-CONTAINING			
DE	PROTEIN) (EVEC).			
GN	FBLN5 OR DANCE.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99357779; PubMed=10428823;			
RA	Nakamura T., Ruiz-Lozano P., Lindner V., Yabe D., Taniwaki M.,			
RA	Furukawa Y., Kobuke K., Tashiro K., Lu Z., Andon N.L., Schaub R.,			
RA	Matsumori A., Sasayama S., Chien K.R., Honjo T.;			
RT	"DANCE, a novel secreted RGD protein expressed in developing,			
RT	atherosclerotic, and balloon-injured arteries."			
RL	J. Biol. Chem. 274:22476-22483(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99278197; PubMed=10347091;			
RA	Kowal R.C., Richardson J.A., Milano J.M., Olson E.N.;			
RT	"EVEC, a novel epidermal growth factor-like repeat-containing protein			
RT	upregulated in embryonic and diseased adult vasculature."			
RL	Circ. Res. 84:1166-1176(1999).			
CC	!- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH			
CC	INTERACTION OF INTEGRINS AND THE RGD MOTIF. COULD BE A VASCULAR			
CC	LIGAND FOR INTEGRIN RECEPTORS AND MAY PLAY A ROLE IN VASCULAR			
CC	DEVELOPMENT AND REMODELING.			
CC	!- SUBCELLULAR LOCATION: SECRETED.			
CC	!- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AF112153; AAD41769.1; -			
DR	EMBL; AF137350; AAD25101.1; -			
DR	InterPro; IPR000152; Asx hydroxyl.			
DR	InterPro; IPR000561; EGF-like.			
DR	InterPro; IPR001881; EGF_Ca.			
DR	Pfam; PF00008; EGF; 4.			
DR	SMART; SM00179; EGF_CA; 4.			
DR	SMART; SM00001; EGF_like; 2.			
DR	PROSITE; PS00010; ASX_HYDROXYL; 4.			
DR	PROSITE; PS00022; EGF_1; FALSE_NEG.			

DE FLAGELLAR BIOSYNTHETIC PROTEIN FLIP.
GN FLIP OR FLAR.
OS Salmonella typhimurium.
OC Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX Salmonella.
RN NCBI_TaxID=602;
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RA Schoenhalz G.J., Ohnishi K., Kihara M., Macnab R.M.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PLAYS A ROLE IN THE FLAGELLUM-SPECIFIC TRANSPORT SYSTEM.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE FLIP/MOPC/SAP FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L49021; AAB81319.1; -
CC StyGene: SG10577; flip.
CC InterPro: IPR002039; Flap.
CC Pfam: PF00813; Flap; 1.
CC PRINTS: PR00951; FLGRIOSNFLP.
CC PRODOM: PD002586; FLIP; 1.
CC PROSITE: PS01060; FLIP_1; 1.
CC PROSITE: PS01061; FLIP_2; 1.
CC Flagella; Transmembrane.
KW TRANSMEM 4 24 POTENTIAL.
FT TRANSMEM 45 65 POTENTIAL.
FT TRANSMEM 88 108 POTENTIAL.
FT TRANSMEM 185 205 POTENTIAL.
FT TRANSMEM 209 229 POTENTIAL.
SQ SEQUENCE 245 AA; 26787 MW; C9A4241F0653A4D4 CRC64;

Query Match 12.6%; Score 76.5; DB 1; Length 245;
Best Local Similarity 30.8%; Pred. No. 1.5;
Matches 33; Conservative 9; Mismatches 38; Indels 27; Gaps 6;

QY 6 LRRLLRLVLGLWALRSVAGEAQP--TAP-CSRGSWSADLDKMCDCASCRARP--- 59
Db 1 MRRLLFLSLAGLW--LFSPAAQAQLGLISQPLAGGGGWSLSVQTLVFTLSLTPAIL 58

QY 60 --HSDF-----CLGCAAPAPFRLLPILGGALSFLVL 92
Db 59 LMMTSFTRIIVFGLLRNALGTSPAPPNQV-----LLGLALFLTFFI 100

RESULT 4
PORD_METH STANDARD; PRT; 81 AA.
AC P56815;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PYRUVATE SYNTHASE SUBUNIT PORD (EC 1.2.7.1) (PYRUVATE OXIDOREDUCTASE
DE DELTA CHAIN) (PORD) (PYRUVIC-FERREDOXIN OXIDOREDUCTASE DELTA SUBUNIT).
GN PORD OR MTH1740.1.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanothermobacter.
OX NCBI_TaxID=145262;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DELTA H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,

RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shmer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997)
CC -!- CATALYTIC ACTIVITY: PYRUVATE + COA + OXIDIZED FERREDOXIN =
CC ACETYL-COA + CO(2) + REDUCED FERREDOXIN.
CC -!- COFACTOR: BINDS TWO 4FE-4S CLUSTERS.
CC -!- SUBUNIT: HETEROTETRAMER OF ONE ALPHA, ONE BETA, ONE DELTA AND ONE
CC GAMMA CHAIN.
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL TYPE FERREDOXIN FAMILY.
CC -!- CAUTION: THERE SEEMS TO BE A SEQUENCING ERROR THAT FUSES TOGETHER
CC PORD AND PORD. WE HAVE CUT THE ORF INTO ITS TWO CONSTITUENTS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE000929; AAB86210.1; ALT_INIT.
CC HSP: P00195; ICLP.
CC InterPro: IPR001450; 4FE4S_ferrdxin.
CC Pfam: PF00037; fer4; 2.
CC PRINTS: PR00353; 4FE4SFERDOXIN.
CC PROSITE: PS00198; 4FE4S_FERREDOXIN; 2.
KW Oxidoreductase; Electron transport; Iron-sulfur; Repeat; 4Fe-4S;
KW Complete proteome.
FT METAL 34 34 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 37 37 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 40 40 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 44 44 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 60 60 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 63 63 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 66 66 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 70 70 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
SQ SEQUENCE 81 AA; 9121 MW; 219A9CCAB8A41604 CRC64;

Query Match 12.1%; Score 73.5; DB 1; Length 81;
Best Local Similarity 29.2%; Pred. No. 1.2;
Matches 19; Conservative 8; Mismatches 21; Indels 17; Gaps 4;

QY 31 PGTAPCSRSWSA-----DLKCMDCASCR-----RPHS---DFCLGCA-AAPPA 73
Db 11 PGSTVKNKTSWRTFRPVLDDKDKCIDCDNCILFCPEGSCINREHIDYDYCKGGICAEC 70

QY 74 PFRL 78
Db 71 PVKAI 75

RESULT 5
NPCI_HUMAN STANDARD; PRT; 1278 AA.
ID NPCI_HUMAN
AC O15118;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE NIEMANN-PICK C1 PROTEIN PRECURSOR.
GN NPC1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANTS NPC1.
RX MEDLINE=97362323; PubMed=9211849;

Matches 26; Conservative 6; Mismatches 19; Indels 18; Gaps 5;

QY 11 RLLVLGLMLA--LLRSVAGEQAPGRAPCSRGSWSADLDKCMDCASC-----ARPHSDFC 64
 Db 221 RSFVQGLGVADWRKVA--QVPLAPECRA-----VMKLVYCAHCRGVPGARPCPDYC 272

QY 65 ----LGCAA 69
 Db 273 RNVLRGGLA 281

RESULT 9
 MLP2_DROME STANDARD; PRT: 495 AA.
 AC Q24400: Q9VI62;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE MUSCLE LIM PROTEIN MLP84B.
 GN LIM3 OR MLP84B OR CGI0699.
 OS Drosophila melanogaster (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96387325; PubMed=8794860;
 RA Stronach B.E., Slegrist S.E., Beckerle M.C.;
 RT "Two muscle-specific LIM proteins in Drosophila.";
 RL J. Cell Biol. 134:11179-1195(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=BERKELEY;
 RC MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
 RA Brudnon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: PLAYS A ROLE IN CELL DIFFERENTIATION LATE IN MYOGENESIS.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR, AND CYTOPLASMIC.
 CC -1- TISSUE SPECIFICITY: IN THE EMBRYO, EXPRESSION IS RESTRICTED TO THE
 CC SOMATIC, VICERAL, AND PHARYNGEAL MUSCLES. WITHIN THE SOMATIC
 CC MUSCULATURE, MLP48B IS LOCALIZED AT THE ENDS OF MUSCLE FIBERS AT
 CC THE POINT OF ATTACHMENT TO THE EPIDERMIS. THERE IS NO EXPRESSION
 CC IN CARDIAC MESODERM OR IN FAT BODY.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSION IS BIPHASIC, PEAKING LATE IN
 CC EMBRYOGENESIS (16-24 H EMBRYOS) AND DURING THE LARVAL TO PUPAL
 CC TRANSITION, WHEN THE MUSCULATURE IS DIFFERENTIATING. FOUND IN
 CC DEVELOPING MUSCLES OF THE VISCERAL AND SOMATIC MESODERM SUBSEQUENT
 CC TO THE FORMATION OF THE MUSCLE PRECURSOR CELLS. DECREASED LEVELS
 CC ARE STILL DETECTABLE IN ADULTS.
 CC -1- SIMILARITY: TO THE VERTEBRATE CYSTEINE-RICH PROTEINS.
 CC -1- SIMILARITY: CONTAINS 5 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
 CC IONS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X91245; CAA62627.1; -;
 DR EMBL; AF090832; AAC61591.1; -;
 DR EMBL; AE003672; AAF54063.1; -;
 DR HSSP; Q05158; IOLI;
 DR FlyBase; FBgn0014863; Mlp84B.
 DR InterPro; IPR001781; LIM.
 DR Pfam; PF00412; LIM; 5.
 DR ProDom; PD000094; LIM; 5.
 DR SMART; SM00132; LIM; 5.
 DR PROSITE; PS00478; LIM_DOMAIN_1; 5.
 DR PROSITE; PS00023; LIM_DOMAIN_2; 5.
 KW Nuclear protein; Repeat; LIM domain; Metal-binding; Zinc; Myogenesis;
 KW Developmental protein; Differentiation.
 FT DOMAIN 12 63
 FT DOMAIN 65 80 GLY-RICH.
 FT DOMAIN 66 71 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 120 172 LIM 2.
 FT DOMAIN 175 180 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 178 189 GLY-RICH.
 FT DOMAIN 222 274 LIM 3.
 FT DOMAIN 276 291 GLY-RICH.
 FT DOMAIN 325 377 LIM 4.
 FT DOMAIN 379 390 GLY-RICH.
 FT DOMAIN 421 473 LIM 5.
 FT DOMAIN 475 490 GLY-RICH.
 SQ SEQUENCE 495 AA; 53525 MW; 2E559B9178E54C0E CRC64;

Query Match 11.4%; Score 69.5; DB 1; Length 495;
 Best Local Similarity 28.0%; Pred. No. 13;
 Matches 23; Conservative 7; Mismatches 37; Indels 15; Gaps 2;

QY 20 ALLRSVAGEAP-----GTAPCSRGSWSADLDKCMDCAS-----CRARPHSDFC 64
 Db 110 AIARAPEGEPCRGVYVAAEQMLARGSNHKEFCGCTCKGLDLSILCCEAPDKNIYC 169
 QY 65 LGCAAPAPPRLLWLPILGGAL 86
 Db 170 KGCYAKFKPGKGYGGGGGAL 191

RESULT 10

```

PTTG_HUMAN
ID PTTG_HUMAN STANDARD; PRT; 180 AA.
AC P53801; Q9NS09;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PITUITARY TUMOR-TRANSFORMING GENE 1 PROTEIN-INTERACTING PROTEIN
DE (PITUITARY TUMOR-TRANSFORMING GENE PROTEIN BINDING FACTOR) (PTTG-
DE BINDING FACTOR) (PBF).
GN PTTGLIP OR C21ORF1 OR C21ORF3.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thymus;
RX MEDLINE=98234553; PubMed=9570958;
RA Yaspo M.-L., Aaltonen J., Horelli-Kuitunen N., Peltonen L.,
RA Lehrach H.;
RT "Cloning of a novel human putative type Ia integral membrane protein
RT mapping to 21q22.3.";
RL Genomics 49:133-136(1998).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordstiek G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Lehrach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21.";
RL Nature 405:311-319(2000).
[3]
RN SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=20317150; PubMed=10781616;
RA Chien W., Pei L.;
RT "A novel binding factor facilitates nuclear translocation and
RT transcriptional activation function of the pituitary
RT tumor-transforming gene product.";
RL J. Biol. Chem. 275:19422-19427(2000).
CC -!- FUNCTION: MAY FACILITATE PTTG NUCLEAR TRANSLOCATION.
CC -!- SUBUNIT: INTERACTS WITH PTTG.
CC -!- SUBCELLULAR LOCATION: ACCORDING TO REF.3, IT IS FOUND IN THE
CC CYTOPLASM AND IN THE NUCLEUS. ACCORDING TO REF.1, IT IS A TYPE I
CC MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: UBIQUITOUS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; 250022; CAA90325.1; -
CC EMBL; AL163300; CAB90552.1; -
CC EMBL; AF149785; AAF73770.1; -
CC MIM; 603784; -
CC InterPro; IPR003659; PSI.
CC SMART; SM00423; PSI; 1.
KW Transmembrane; Nuclear protein.
FT TRANSMEM 97 117 POTENTIAL.
FT DOMAIN 116 121 POLY-CYS.
FT CONFLICT 9 19 PTFYWRRLRG --> ARRTGGCASV (IN REF. 3).

```

```

FT CONFLICT 108 114 TLLLGIA -> NPPPGHC (IN REF. 3).
SQ SEQUENCE 180 AA; 20323 MW; F1E66014D49EC1DE CRC64;

Query Match 11.3%; Score 69; DB 1; Length 180;
Best Local Similarity 25.8%; Pred. No. 6.1;
Matches 32; Conservative 19; Mismatches 43; Indels 30; Gaps 7;

QY 10 LRLVLGLWLLRSVAGEQAPGTAPCSRSGSSWSADLDKCMDCASCRCRPHSDFCILG---66
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 15 LRLGGAULLLLIPVAAAEPPGAA--CSQNTKTC--EECLKNVSCIMGNTNKCACLDYPV 71
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 67 CAAPAP-----FRLW-----PILGALSILTFVLGLSLGFLVWRRCRRR 108
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 72 TSVLPASLCKLSSARWGVWVNFALITMSVGG-----TLLLGI--AICCCCCRRR 125
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 109 SSPP 112
   |
DB 136 SRKP 129

RESULT 11
Y127_HUMAN
ID Y127_HUMAN STANDARD; PRT; 314 AA.
AC Q14140;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36; Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN KIAA0127.
GN KIAA0127.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP TISSUE=Bone marrow;
RC SEQUENCE FROM N.A.
RX MEDLINE=96127530; PubMed=8590280;
RA Nagase T., Seki N., Tanaka A., Ishikawa K.-I., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. IV.
RT The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by
RT analysis of cDNA clones from human cell line KG-1.";
RL DNA Res. 2:167-174(1995).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D50917; BAA09476.1; -
CC KW Hypothetical protein.
SQ SEQUENCE 314 AA; 33896 MW; E43107FC565AAC31 CRC64;

Query Match 11.3%; Score 69; DB 1; Length 314;
Best Local Similarity 34.0%; Pred. No. 9.7;
Matches 18; Conservative 6; Mismatches 27; Indels 2; Gaps 1;

QY 30 APGTAPCSRSGSSWSADLDKCMDCASCRCRPHSDFCILGCAAPAPPRLLWPIL 82
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 112 SPSSHPCDLGS--TTPLACLTPASILLEDDDDTCTCSQAMQPTAPTKLSPPAL 162
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
FASA_RAT
ID FASA_RAT STANDARD; PRT; 324 AA.
AC Q63199;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38; Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)

```

DE FASL RECEPTOR PRECURSOR (APOPTOSIS-MEDIATING SURFACE ANTIGEN FAS)
 DE (APO-1 ANTIGEN) (CD95).
 GN TNFRSF6 OR PTI OR FAS.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;
 RX MEDLINE=94128114; PubMed=7507668;
 RA Kimura K., Yamamoto M., Wakatsuki T.;
 RT "A variant mRNA species encoding a truncated form of Fas antigen in
 RL the rat liver.";
 RL Biochem. Biophys. Res. Commun. 198:666-674(1994).
 CC -1- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS FASL. THE
 CC ADAPTOR MOLECULE FADD RECRUITS CASPASE-8 TO THE ACTIVATED
 CC RECEPTOR. THE RESULTING AGGREGATE CALLED THE DEATH-INDUCING
 CC SIGNALING COMPLEX (DISC) PERFORMS CASPASE-8 PROTEOLYTIC
 CC ACTIVATION. ACTIVE CASPASE-8 INITIATES THE SUBSEQUENT CASCADE OF
 CC CASPASES (ASPARTATE-SPECIFIC CYSTEINE PROTEASES) MEDIATING
 CC APOPTOSIS. FAS-MEDIATED APOPTOSIS MAY HAVE A ROLE IN THE
 CC INDUCTION OF PERIPHERAL TOLERANCE. IN THE ANTIGEN-STIMULATED
 CC SUICIDE OF MATURE T-CELLS, OR BOTH (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- DOMAIN: CONTAINS A DEATH DOMAIN INVOLVED IN THE BINDING OF FADD,
 CC AND MAYBE TO OTHER CYTOSOLIC ADAPTOR PROTEINS.
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D26112; BAA05108.1; -;
 DR InterPro; IPR000488; Death.
 DR InterPro; IPR001368; TNFR_c6.
 DR Pfam; PF00531; death; 1.
 DR Pfam; PF00020; TNFR_c6; 3.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00208; TNFR; 3.
 DR PROSITE; PS00652; TNFR_NGFR_1; 2.
 DR PROSITE; PS50050; TNFR_NGFR_2; 2.
 DR PROSITE; PS50017; DEATH_DOMAIN; 1.
 KW Apoptosis; Receptor; Glycoprotein; Transmembrane; Repeat; Signal.
 FT SIGNAL 1 21 BY SIMILARITY.
 FT CHAIN 22 324 FASL RECEPTOR.
 FT DOMAIN 22 171 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 172 188 POTENTIAL.
 FT DOMAIN 189 324 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 43 163 3 X TNFR-CYS.
 FT REPEAT 43 79 TNFR-CYS 1.
 FT REPEAT 80 123 TNFR-CYS 2.
 FT REPEAT 124 163 TNFR-CYS 3.
 FT DOMAIN 219 303 DEATH.
 FT CARBOHYD 43 43 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 132 132 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 324 AA; 36835 MW; D25D583C909D9D09 CRC64;

Query Match 11.3%; Score 69; DB 1; Length 324;
 Best Local Similarity 18.2%; Pred. No. 10;
 Matches 24; Conservative 18; Mismatches 34; Indels 56; Gaps 5;
 Qy 30 APGTPAPCRSGWSAD---LDRKMCDA-----SCRARPH----- 60
 || || | : : ||| ||
 Db 75 APTCHPCTEGEYTDKHKYSKRCRCACAFCDGHLGVETNCTRTQNTKCRCKENFCNAS 134
 || | : : ||| ||

QY 61 -SDFCLCAAA-----PPAFRLLPILGGALSITFVLGLLSGLV 100
 | | : : ||| : : | : | :
 Db 135 LCDHCYCHTSCGLEIDLEPCTRTSNTCKKSSNYKLLWLLPLGLALFV-----FI 187
 QY 101 WRCRRERSPP 112
 Db 188 YKRYKRQPGDP 199
 RESULT 13
 GSC_L_HUMAN STANDARD; PRT; 205 AA.
 ID GSC_L_HUMAN AC O15499;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE 30-MAY-2000 (Rel. 39, Last annotation update)
 DE HOMEBOX PROTEIN GOOSECOID-LIKE (GSC-2).
 GN GSC_L.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANT CYS-47.
 RX MEDLINE=98110571; PubMed=9441739;
 RA Funke B., St Jore B., Puech A., Sirotkin H., Edelman L., Carlson C.,
 RA Raft S., Pandita R.K., Kucherlapati R., Skoultschi A., Morrow B.E.;
 RT "Characterization and mutation analysis of gooseoid-like (GSL), a
 RT homeodomain-containing gene that maps to the critical region for
 RT VCFS/DGS on 22q11.";
 RL Genomics 46:364-372(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97294411; PubMed=9150167;
 RA Gottlieb S., Emanuel B.S., Driscoll D.A., Sellinger B., Wang Z.,
 RA Roe B., Budarf M.L.;
 RT "The DiGeorge syndrome minimal critical region contains a gooseoid-
 RT like (GSL) homeobox gene that is expressed early in human
 RT development.";
 RL Am. J. Hum. Genet. 60:1194-1201(1997).
 CC -1- FUNCTION: MAY HAVE A ROLE IN DEVELOPMENT. MAY REGULATE ITS OWN
 CC TRANSCRIPTION. MAY BIND THE BICOID CONSENSUS SEQUENCE TAATCC.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: DETECTED IN ADULT TESTIS AND PITUITARY, AND IN
 CC 9-10 WK FETAL TISSUE (THORAX). PROBABLY EXPRESSED IN OTHER TISSUES
 CC AT LOW LEVELS.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN EARLY HUMAN DEVELOPMENT AS WELL
 CC AS IN A LIMITED NUMBER OF ADULT TISSUES.
 CC -1- DISEASE: MAY BE RELATED TO SOME OF THE DEVELOPMENTAL DEFECTS
 CC ASSOCIATED WITH VELOCARDIOFACIAL AND DIGEORGE SYNDROMES
 CC (VCFS/DGS).
 CC -1- SIMILARITY: BELONGS TO THE PAIRED FAMILY OF HOMEBOX PROTEINS.
 CC "BICOID" SUBFAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U96402; AAC39544.1; -;
 DR HSSP; P06601; IFJL.
 DR TIM; 601845; -;
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.
 DR PROSITE; PS50071; HOMEBOX_2; 1.
 KW Homeobox; DNA-binding; Nuclear protein; Polymorphism.
 FT DOMAIN 2 5 POLY-ALA.
 FT DOMAIN 64 70 POLY-CYS.

Search completed: March 12, 2002, 10:00:06
Job time: 188 sec